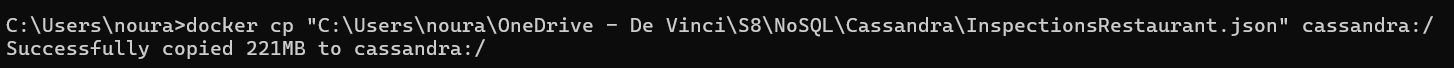
# TP Cassandra :

## Chapter 1 : Create a database on Cassandra



Une image contenant texte, capture d’écran, Police

Description générée automatiquement

### Creating the tables :

Une image contenant texte, capture d’écran, Police

Description générée automatiquement



While trying to create the inspections table using a foreign key for idRestaurant, we found out that Cassandra doesn’t use foreign keys.Une image contenant texte, capture d’écran, Police

Description générée automatiquement

**Why No Foreign Keys in Cassandra?**

Cassandra avoids enforcing foreign key constraints because:

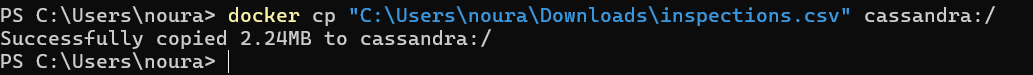
1. **Distributed Architecture**: Verifying foreign key constraints in a distributed system would require cross-node communication, which can be slow and inefficient.
2. **Denormalized Data**: Cassandra encourages duplicating data (denormalization) to optimize for read performance rather than maintaining relational integrity

Une image contenant texte, capture d’écran, Police

Description générée automatiquement



### Adding the data to cassandra :



### Loading the data into the tables :

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

### SIMPLE QUERIES :

Query 1 : List all the restaurants.

SELECT \* FROM RESTAURANTS ;

Une image contenant capture d’écran, texte, Logiciel multimédia, logiciel

Description générée automatiquement

Query 2 : get inspection details from a specific restaurant.

SELECT \* FROM inspections WHERE idRestaurant = 50016731 ;

Une image contenant capture d’écran, texte, Police

Description générée automatiquement

Query 3 : count the total number of inspections.

SELECT COUNT(\*) FROM inspections;

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Query 4 : get the average inspection score.

SELECT AVG(score) FROM inspections ;

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Query 5 : count the restaurants that are in Manhattan

SELECT COUNT(\*) FROM restaurants WHERE borough = ‘MANHATTAN’;

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Query 6 : get the maximum score of an inspection.

SELECT MAX(score) FROM inspections;

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

## Complex Queries :

Query 1 : Get the number of inspections per restaurant

SELECT idRestaurant, COUNT(\*) FROM inspections GROUP BY idRestaurant;

Une image contenant texte, capture d’écran, logiciel, Logiciel multimédia

Description générée automatiquement

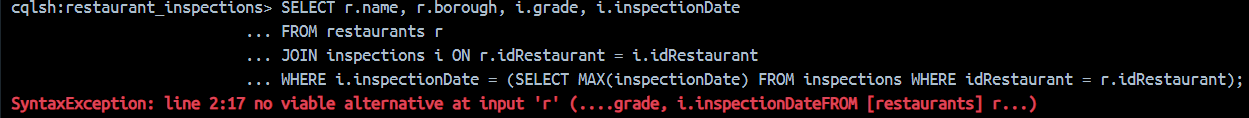
Query 2 : get the average score of inspections per restaurant.

Une image contenant texte, capture d’écran, logiciel

Description générée automatiquement

### Hard Queries :

Get restaurant details with their inspection date.



CQL does not support JOINS, instead, we use denormalization.

Step 1 : Create a denormalized Table

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Step 2 : Preprocess the data and load it to cassandra.

We used python to generate a new csv file only containing the columns that are needed.



Une image contenant texte, capture d’écran, Police

Description générée automatiquement

We should not worry about the error message as it is only written because Cassandra tried to insert the header line.

Step 3 : Write the query and see the result.

Une image contenant texte, capture d’écran

Description générée automatiquement