

TP06: NoSQL Databases – Apache Cassandra

kiram mohammed ali

Introduction

This report presents the solution to TP06: NoSQL Databases using Apache Cassandra.¹ The goal is to create a keyspace, import restaurant and inspection data, and execute a series of CQL queries to explore the dataset. All screenshots are stored in the folder `photo/`.

1. Listing all Restaurants

Query

```
SELECT * FROM Restaurant;
```

Screenshot

id	borough	buildingid	cuisine_type	phone	street	zipcode	name	hours	price	type
40000001	STATEN ISLAND	1545								
40000002	QUEENS	RUSTEN MARIOTT	7183511508		FOREST AVENUE	11382	American			Pizza
40000003	QUEENS	15100					American			
40000004	QUEENS	LENDIN'S RESTAURANT	7183469778		ATLANTIC AVENUE	11418	American			Pizzeria
40000005	MANHATTAN	1460								
40000006	MANHATTAN	BANQUO HOUSE	2125515943		WEST 46 STREET	10018	Caribbean			
40000007	MANHATTAN	2225								
40000008	MANHATTAN	7175 TASTY CORNER	7183448473		TILDEN AVENUE	11226	American			
40000009	QUEENS	1224								
40000010	QUEENS	WATSON'S HOT HOUSE	7183595038		ROOSEVELT AVENUE	11368	Chinese			
40000011	STATEN ISLAND	77								
40000012	QUEENS	YIMMY YU'S	7183423888		PORT MICHIGAN AVE	11382	Chinese			
40000013	QUEENS	4000								
40000014	QUEENS	KITCHEN	7183551388		PORT MICHIGAN AVE	11329	Italian			
40000015	MANHATTAN	157								
40000016	MANHATTAN	276								
40000017	MANHATTAN	332								
40000018	MANHATTAN	435								
40000019	MANHATTAN	435								
40000020	MANHATTAN	435								
40000021	MANHATTAN	435								
40000022	MANHATTAN	435								
40000023	MANHATTAN	435								
40000024	MANHATTAN	435								
40000025	MANHATTAN	435								
40000026	MANHATTAN	435								
40000027	MANHATTAN	435								
40000028	MANHATTAN	435								
40000029	MANHATTAN	435								
40000030	MANHATTAN	435								
40000031	MANHATTAN	435								
40000032	MANHATTAN	435								
40000033	MANHATTAN	435								
40000034	MANHATTAN	435								
40000035	MANHATTAN	435								
40000036	MANHATTAN	435								
40000037	MANHATTAN	435								
40000038	MANHATTAN	435								
40000039	MANHATTAN	435								
40000040	MANHATTAN	435								
40000041	MANHATTAN	435								
40000042	MANHATTAN	435								
40000043	MANHATTAN	435								
40000044	MANHATTAN	435								
40000045	MANHATTAN	435								
40000046	MANHATTAN	435								
40000047	MANHATTAN	435								
40000048	MANHATTAN	435								
40000049	MANHATTAN	435								
40000050	MANHATTAN	435								
40000051	MANHATTAN	435								
40000052	MANHATTAN	435								
40000053	MANHATTAN	435								
40000054	MANHATTAN	435								
40000055	MANHATTAN	435								
40000056	MANHATTAN	435								

Figure 1: Result of SELECT * FROM Restaurant

¹Based on the official TP instructions.

Overview

This query returns all rows from the `Restaurant` table (about 25,624 restaurants imported from CSV).

2. Listing all Restaurant Names

Query

```
SELECT name FROM Restaurant;
```



```
cqlsh:resto_ny> SELECT name FROM Restaurant;
name
-----
BOSTON MARKET
LENIHAN'S SALOON
BANGKOK HOUSE
TJ'S TASTY CORNER
NATHAN'S HOT DOGS
YUMMY YUMMY
KING'S KITCHEN
CESCA
EL GUANACO RESTAURANT & PUPUSERIA
THE BROOK
LEO'S DELI & GRILL
KAEDE JAPANESE RESTAURANT
JOHNS PIZZA
RICE & BEANS LECHONERA
INDIAN EXPRESS
DOUBLETREE GREENHOUSE 36
VIDAELVA
TK VILLAGE BAKERY COMPANY
BKB
SPRING
DUNKIN' DONUTS
PLANET WINGS
NEW CHINA RESTAURANT WANG INC
JIN SUSHI & THAI
JADE ASIAN EXPRESS INC
TIELLA RESTAURANT
CHOLE'S RESTAURANT & LOUNGE
SUNDAES AND CONES
ON 14TH STREET
STARBUCKS
HALE & HEARTY SOUP
MURRAY'S CHEESE BAR
GOLDEN KRUST
EATING TREE
DRINK
BISOUX MARKET/WORLD BEAN
TO YOUR HEALTH & HAPPINESS
P.J.' S
PINKBERRY
KING OF TANDOOR
PRONTO DELI
FLUSHING HOUSE RESIDENCE FOR ADULTS (SENIOR HOME)
EL PALENQUE SPORT BAR
RED ROSE
WANG EASTERN
AROY DEE THAI KITCHEN
SUBWAY
TSOB-TSOBE CAFE LOUNGE BAR
HEALTHY HENRY'S
STARBUCKS
(25624 rows)
```

Figure 2: List of restaurant names

Overview

This returns only the `name` column for all restaurants.

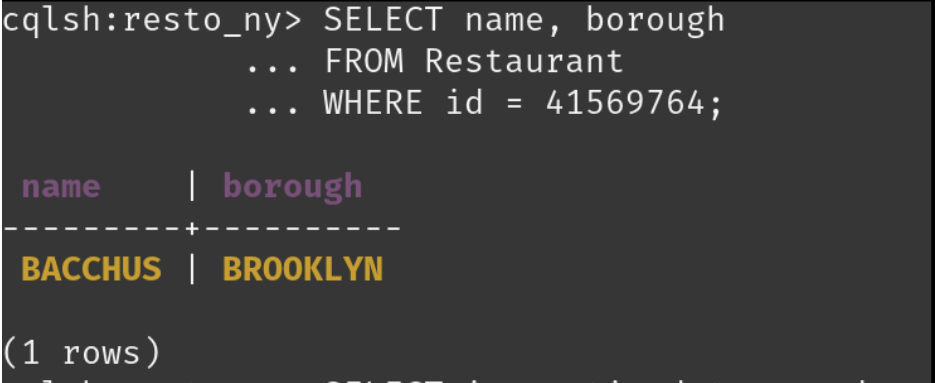
3. Name and Borough of Restaurant ID = 41569764

Query

```
SELECT name, borough
```

```
FROM Restaurant
WHERE id = 41569764;
```

Screenshot



A screenshot of a terminal window with a dark background. The prompt is 'cqlsh:resto_ny>'. The query entered is 'SELECT name, borough ... FROM Restaurant ... WHERE id = 41569764;'. The result is displayed as a table with two columns: 'name' and 'borough'. The first row contains the values 'BACCHUS' and 'BROOKLYN'. Below the table, it says '(1 rows)'.

```
cqlsh:resto_ny> SELECT name, borough
... FROM Restaurant
... WHERE id = 41569764;

name | borough
-----+-----
BACCHUS | BROOKLYN

(1 rows)
```

Figure 3: Restaurant with ID = 41569764

Overview

The restaurant with this ID is **BACCHUS** located in **BROOKLYN**.

4. Inspection Dates and Grades for Restaurant 41569764

Query

```
SELECT inspectiondate, grade
FROM Inspection
WHERE idrestaurant = 41569764;
```

Screenshot

```
cqlsh:resto_ny> SELECT name, borough
... FROM Restaurant
... WHERE id = 41569764;
```

name	borough
BACCHUS	BROOKLYN

(1 rows)

```
cqlsh:resto_ny> SELECT inspectiondate, grade
... FROM Inspection
... WHERE idrestaurant = 41569764;
```

inspectiondate	grade
2013-06-27	null
2013-07-08	A
2013-12-26	null
2014-02-05	A
2014-07-17	null
2014-08-06	A
2015-01-08	A
2016-02-25	A

(8 rows)

```
cqlsh:resto_ny> 
```

Figure 4: Inspection history for restaurant 41569764

Overview

A total of 8 inspections were returned, with grades alternating between A and null.

5. Restaurants Serving French Cuisine

Query

```
SELECT name
FROM Restaurant
WHERE cuisinetype = 'French';
```



```
cqlsh:resto_ny> SELECT name
... FROM Restaurant
... WHERE cuisinetype = 'French';

name
-----
MATISSE
ALMANAC
TOUT VA BIEN
FELIX
CREPES ON COLUMBUS
THE BARONESS BAR
THE SIMONE
FP BAKERY
VIN ET FLEURS
CAFE BOULUD/BAR PLEIADES
COCOTTE
Bourgeois Pig
DELICE & SARRASIN
LA TARTE FLAMBEE
JEAN GEORGES
MAISON MAY
DANIEL
SAJU BISTRO
LE PAIN QUOTIDIEN
CAFE CLUNY
BIN 71
CREPES CELSTES
JOYCE BAKESHOP
THE FOX AND THE CREPES

MATCH 65
AIR FRANCE LOUNGE
THE TEN BELLS
BRASSERIE MAGRITTE
PAPILLON BISTRO & BAR
LE GARAGE
CREDIT AGRICOLE-CBI EXECUTIVE ROOM
RAOULS
LA CREPE C'EST SI BON
VIVE LA CREPE
THE PARK BENCH CAFE & CREPERIE
LE PAIN QUOTIDIEN
OCABANON
CAFE CHAMPIGNON
PERRY STREET
THE BROOKLYN LABEL
LE BILBOQUET
BRASSERIE COGNAC
TALON BAR
FINANCIER PATISSERIE
BOUCHON BAKERY
LE GIGOT
THE PANDERING PIG
TARTINERY
SANTOS ANNE
CAFE DADA
BAR SUZETTE CREPERIE

(351 rows)
cqlsh:resto_ny>
```

Figure 5: French restaurants

Overview

351 restaurants serve French cuisine.

6. Restaurants in BROOKLYN

Query

```
SELECT name
FROM Restaurant
WHERE borough = 'BROOKLYN';
```

Note

If Cassandra rejects the query, create an index:

```
CREATE INDEX resto_borough_idx ON Restaurant (borough);
```

```
cqlsh:resto_ny> CREATE INDEX resto_borough_idx ON Restaurant (borough);
cqlsh:resto_ny> SELECT name
... FROM Restaurant
... WHERE borough = 'BROOKLYN';

name
-----
TP'S TASTY CORBEN
KING'S KITCHEN
LEO'S DELI & GRILL
JIN SUSHI & THAI
CROWN FRIED CHICKEN
BROOKLYN CAFE 1
CRESCENT COFFEE SHOP
LA ROYAL'S FRIED BURGERS HOUSE
CONNECTICUT MUFFIN
GREENSTREET'S SALADS
THE TOPAZ
LA GUSTAZIA SAN
MAMA RING SOUL FOOD
HONG HONG BAKERY
BROOKLYN BRIDGE GARDEN BAR
THE CANTINE
THE GUMBO SHED
DAVEY'S ICE CREAM
CROWN FRIED CHICKEN
THAI TONY'S
CHINA DRAGON
BROADWAY BAGEL CAFE
```

```
SUBWAY
NIGHTS AND WEEKENDS BK
EAGLE CITY TEX MEX CHINESE FOOD
DELI & GRILL
LAS CONCHITAS BAKERY
PALO SANTO
SHANG HAI FAMILY DUMPLING
STARLINER CAFE AND MARKET
YEMEN CAFE & RESTAURANT
LITTLE ITALY M & O PIZZA RESTAURANT
RECONNECT CAFE
JOE'S PIZZA
LUIGI'S PIZZA II
CONSTANTINE RESTAURANT AND BAR
FLOYD, NY
PARKSIDE COFFEE SHOP
IKEA
MABLE'S SMOKEHOUSE AND BANQUET HALL
RED MANGO
NEW TOPPS BAKERY & RESTAURANT
SMOKESHACK
GEORGIAN DELI AND BAKERY
CHOLE'S RESTAURANT & LOUNGE
GOLDEN KRUST
KING OF TANDOOR
RED ROSE
TSOB-TSOBE CAFE LOUNGE BAR

(6260 rows)
cqlsh:resto_ny>
```

Figure 6: Restaurants located in BROOKLYN

Overview

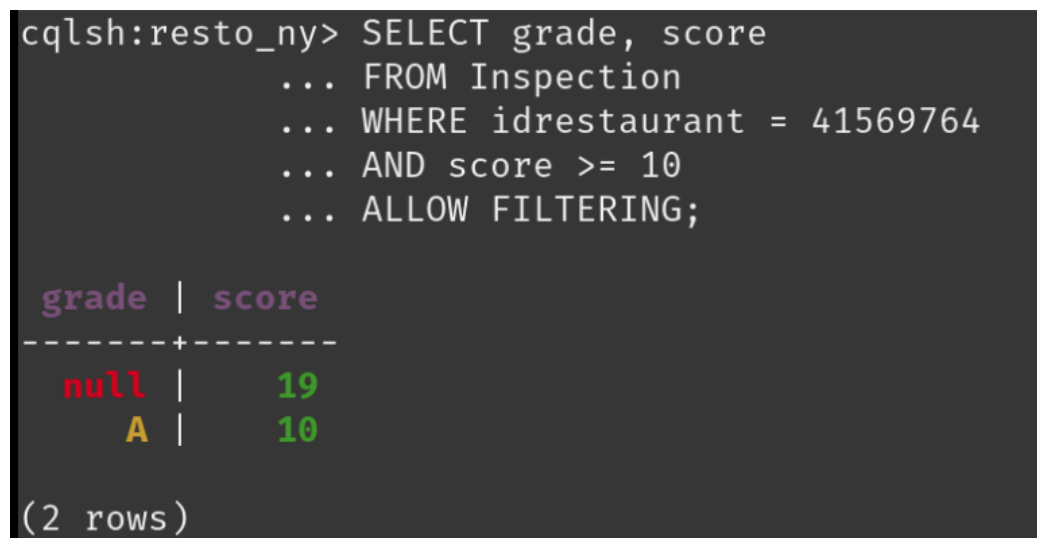
This query returns about 6,260 restaurants in Brooklyn.

7. Grades and Scores ≥ 10 for Restaurant 41569764

Query

```
SELECT grade, score
FROM Inspection
WHERE idrestaurant = 41569764
AND score >= 10
ALLOW FILTERING;
```

Screenshot



The screenshot shows a terminal window with the following text:

```
cqlsh:resto_ny> SELECT grade, score
... FROM Inspection
... WHERE idrestaurant = 41569764
... AND score >= 10
... ALLOW FILTERING;
```

grade	score
null	19
A	10

(2 rows)

Figure 7: Grades and scores ≥ 10

Overview

Two results match the condition: score ≥ 10 and score = 19.

8. Non-null Grades with Score > 30

Query

```
SELECT grade
```

```
FROM Inspection
WHERE score > 30 AND grade > ' '
ALLOW FILTERING;
```

Screenshot

```
cqlsh:resto_ny> SELECT grade, score
... FROM Inspection
... WHERE idrestaurant = 41569764
... AND score >= 10
... ALLOW FILTERING;

grade | score
-----+-----
null  | 19
A     | 10

(2 rows)
cqlsh:resto_ny> SELECT grade
... FROM Inspection
... WHERE score > 30 AND grade > ''
... ALLOW FILTERING;

grade
-----
C
C
C
C
Not Yet Graded
C
C
C
C
C
C
C

(1152 rows)

Warnings :
Read 7436 live rows and 5203 tombstone cells for
urant) >= token(50018028) AND grade > '' AND sco
tone_warn_threshold)

cqlsh:resto_ny> 
```

Figure 8: Non-null grades for inspections with score > 30

Overview

1152 rows contain a non-null grade and a score above 30. The query triggers warnings due to large scans and tombstones.

9. Counting Rows from Previous Query

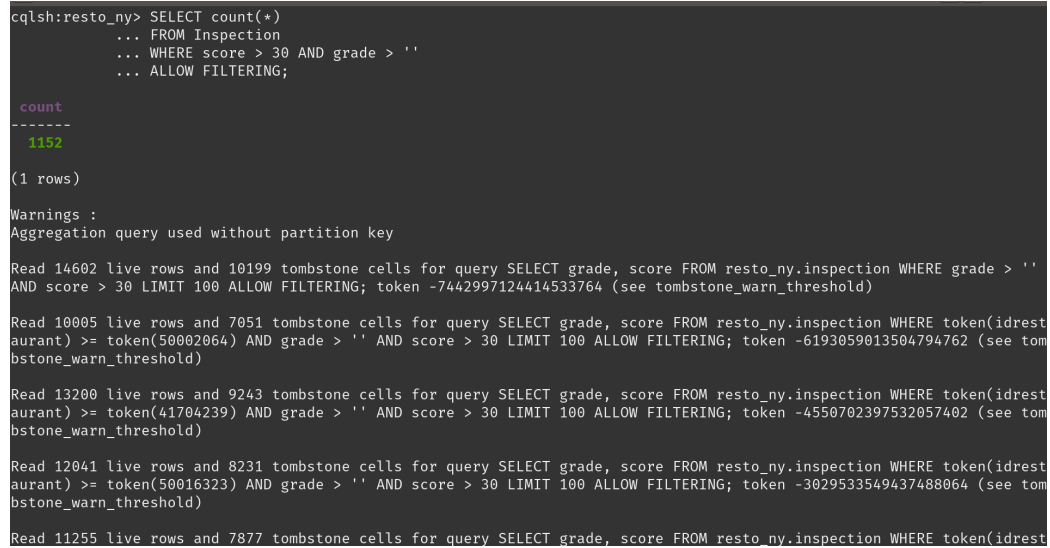
Query

```
SELECT count(*)
FROM Inspection
```



```
WHERE score > 30 AND grade > ''  
ALLOW FILTERING;
```

Screenshot



```
cqlsh:resto_ny> SELECT count(*)  
... FROM Inspection  
... WHERE score > 30 AND grade > ''  
... ALLOW FILTERING;  
  
count  
-----  
1152  
  
(1 rows)  
  
Warnings :  
Aggregation query used without partition key  
  
Read 14602 live rows and 10199 tombstone cells for query SELECT grade, score FROM resto_ny.inspection WHERE grade > ''  
AND score > 30 LIMIT 100 ALLOW FILTERING; token -7442997124414533764 (see tombstone_warn_threshold)  
  
Read 10005 live rows and 7051 tombstone cells for query SELECT grade, score FROM resto_ny.inspection WHERE token(idrest  
aurant) >= token(50002064) AND grade > '' AND score > 30 LIMIT 100 ALLOW FILTERING; token -6193059013504794762 (see tom  
bstone_warn_threshold)  
  
Read 13200 live rows and 9243 tombstone cells for query SELECT grade, score FROM resto_ny.inspection WHERE token(idrest  
aurant) >= token(41704239) AND grade > '' AND score > 30 LIMIT 100 ALLOW FILTERING; token -4550702397532057402 (see tom  
bstone_warn_threshold)  
  
Read 12041 live rows and 8231 tombstone cells for query SELECT grade, score FROM resto_ny.inspection WHERE token(idrest  
aurant) >= token(50016323) AND grade > '' AND score > 30 LIMIT 100 ALLOW FILTERING; token -3029533549437488064 (see tom  
bstone_warn_threshold)  
  
Read 11255 live rows and 7877 tombstone cells for query SELECT grade, score FROM resto_ny.inspection WHERE token(idrest
```

Figure 9: Counting matching rows (1152 total)

Overview

The count confirms that **1152 inspections** match the criteria.

Conclusion

This TP demonstrated the use of Cassandra for large-scale storage: creating a keyspace, generating tables, importing CSV data, and performing CQL queries. Some queries required indexes or `ALLOW FILTERING` due to Cassandra's distributed nature.