# Oracle® NoSQL Database SQL Beginner's Guide





Oracle NoSQL Database SQL Beginner's Guide, Release 22.2

E85380-22

Copyright © 2011, 2022, Oracle and/or its affiliates.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs) and Oracle computer documentation or other Oracle data delivered to or accessed by U.S. Government end users are "commercial computer software" or "commercial computer software documentation" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, reproduction, duplication, release, display, disclosure, modification, preparation of derivative works, and/or adaptation of i) Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs), ii) Oracle computer documentation and/or iii) other Oracle data, is subject to the rights and limitations specified in the license contained in the applicable contract. The terms governing the U.S. Government's use of Oracle cloud services are defined by the applicable contract for such services. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle, Java, and MySQL are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Inside are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Epyc, and the AMD logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

# Contents

Conventions Used in This Book	Vi
Introduction to SQL for Oracle NoSQL Databa	se
Simple SELECT Queries	
SQLBasicExamples Script	2-1
Starting the SQL Shell Choosing column data	2-2 2-2
Substituting column names for a query	2-3
Computing values for new columns	2-3
Identifying tables and their columns	2-4
Filtering Results	2-5
Grouping Results	2-7
Ordering Results	2-7
Limiting and Offsetting Results	2-9
Using External Variables	2-10
Working with complex data	
SQLAdvancedExamples Script	3-1
Working with Timestamps	3-4
Working With Arrays	3-5
Working with Records	3-11
Using ORDER BY to Sort Results	3-12
Working With Maps	3-13
Using the size() Function	3-16
Working with JSON	
SQLJSONExamples Script	4-1



Basic Queries	4-4
Using WHERE EXISTS with JSON	4-5
Seeking NULLS in Arrays	4-7
Examining Data Types JSON Columns	4-8
Using Map Steps with JSON Data	4-11
Casting Datatypes	4-12
Using Searched Case	4-13
Working With GeoJSON Data	
Geodetic Coordinates	5-1
GeoJSON Data Definitions	5-2
Searching GeoJSON Data	5-5
Working With Indexes	
Basic Indexing	6-1
Using Index Hints	6-2
Complex Indexes	6-3
Multi-Key Indexes	6-4
Working with Table Rows	
Adding Table Rows using INSERT and UPSERT	7-1
Modifying Table Rows using UPDATE Statements	7-4
Example Data	7-4
Changing Field Values	7-4
Modifying Array Values	7-6
Adding Elements to an Array	7-6
Changing an Existing Element in an Array	7-9
Removing Elements from Arrays	7-9
Modifying Map Values	7-12
Removing Elements from a Map	7-13
Adding Elements to a Map	7-14
Updating Existing Map Elements	7-16
Managing Time to Live Values	7-20
Avoiding the Read-Modify-Write Cycle	7-22
Introduction to the SQL for Oracle NoSQL Database Shell	
Running the SQL Shell	A-1



Configuring the shell	A-2
Shell Utility Commands	A-3
connect	A-3
consistency	A-3
describe	A-4
durability	A-5
exit	A-5
help	A-6
history	A-6
import	A-6
load	A-7
mode	A-7
output	A-11
page	A-11
show faults	A-11
show indexes	A-12
show mrtable-agent-statistics	A-12
show namespaces	A-19
show query	A-19
show regions	A-20
show roles	A-20
show tables	A-20
show users	A-22
timeout	A-22
timer	A-23
verbose	A-23
version	A-23



#### **Preface**

This document is intended to provide a rapid introduction to the SQL for Oracle NoSQL Database and related concepts. SQL for Oracle NoSQL Database is an easy to use SQL-like language that supports read-only queries and data definition (DDL) statements. This document focuses on the query part of the language. For a more detailed description of the language (both DDL and query statements), see *SQL Reference Guide*.

This book is aimed at developers who are looking to manipulate Oracle NoSQL Database data using a SQL-like query language. Knowledge of standard SQL is not required but it does allow you to easily learn SQL for Oracle NoSQL Database.

#### Conventions Used in This Book

The following typographical conventions are used within this manual:

Information that you are to type literally is presented in monospaced font.

Variable or non-literal text is presented in *italics*. For example: "Go to your *KVHOME* directory."

Case-insensitive keywords, like SELECT, FROM, WHERE, ORDER BY, are presented in UPPERCASE.

Case sensitive keywords, like the function size(item) are presented in lowercase.



Finally, notes of special interest are represented using a note block such as this.



1

# Introduction to SQL for Oracle NoSQL Database

Welcome to SQL for Oracle NoSQL Database. This language provides a SQL-like interface to Oracle NoSQL Database that can be used from a command line interface, scripts, or from the Oracle NoSQL Database Java Table Driver. The SQL for Oracle NoSQL Database data model supports flat relational data, hierarchical typed (schema-full) data, and schema-less JSON data. SQL for Oracle NoSQL Database is designed to handle all such data in a seamless fashion without any "impedance mismatch" among the different sub models.

For information on the command line shell you can use to run SQL for Oracle NoSQL Database queries, see Introduction to the SQL for Oracle NoSQL Database Shell. For information on executing SQL queries from the Oracle NoSQL Database Java Table Driver, see Java Direct Driver Developer's Guide .



# Simple SELECT Queries

This section presents examples of simple queries for relational data. To follow along with the examples, get the <code>Examples</code> download from here and run the <code>SQLBasicExamples</code> script found in the <code>sql</code> folder. The script creates the table as shown, and imports the data.

### SQLBasicExamples Script

The script SQLBasicExamples creates the following table:

```
CREATE TABLE Users (
  id integer,
  firstname string,
  lastname string,
  age integer,
  income integer,
  primary key (id)
);
```

The script also load data into the Users table with the following rows (shown here in JSON format):

```
"id":1,
"firstname": "David",
"lastname": "Morrison",
"age":25,
"income":100000,
"id":2,
"firstname": "John",
"lastname": "Anderson",
"age":35,
"income":100000,
"id":3,
"firstname": "John",
"lastname": "Morgan",
"age":38,
"income":null,
"id":4,
```

```
"firstname":"Peter",
    "lastname":"Smith",
    "age":38,
    "income":80000,
}

{
    "id":5,
    "firstname":"Dana",
    "lastname":"Scully",
    "age":47,
    "income":400000,
}
```

You run the SQLBasicExamples script using the load command:

```
> cd <installdir>/examples/sql
> java -jar <KVHOME>/lib/sql.jar -helper-hosts <host>:<port> \
-store <storename> load \
-file <KVHOME>/examples/sql/SQLBasicExamples.cli
```

#### Starting the SQL Shell

You can run SQL queries and execute DDL statements directly from the SQL shell. This is described in Introduction to the SQL for Oracle NoSQL Database Shell. To run the queries shown in this document, start the SQL shell as follows:

```
java -jar KVHOME/lib/sql.jar
-helper-hosts node01:5000 -store kvstore
sql->
```



This document shows examples displayed in COLUMN mode, although the default output type is JSON. Use the <code>mode</code> command to toggle between COLUMN and JSON (or JSON pretty) output.

#### Choosing column data

You can choose columns from a table. To do so, list the names of the desired table columns after SELECT in the statement, before noting the table after the FROM clause.

The FROM clause can name only one table. To retrieve data from a child table, use dot notation, such as parent.child.

To choose all table columns, use the asterisk (\*) wildcard character as follows:

```
sql-> SELECT * FROM Users;
```



#### The SELECT statement displays these results:

id	firstname	Ì	lastname	İ	age	Ì	income	İ
3     4     2     5	John Peter John Dana David		Morgan Smith Anderson Scully Morrison	1 1 1 1	38 38 35 47 25		NULL 80000 100000 400000 100000	1 1 1 1

5 rows returned

To choose specific column(s) from the table Users, include the column names as a commaseparated list in the SELECT statement:

5 rows returned

5 rows returned

### Substituting column names for a query

You can use a different name for a column during a SELECT statement. Substituting a name in a query does not change the column name, but uses the substitute in the returned data returned. In the next example, the query substitutes Surname for the actual column name lastname, by using the actual-name AS substitute-name clause, in the SELECT statement.

```
sql-> SELECT lastname AS Surname FROM Users;
+-----+
| Surname |
+----+
| Scully |
| Smith |
| Morgan |
| Anderson |
| Morrison |
+-----+
```



#### Computing values for new columns

The SELECT statement can contain computational expressions based on the values of existing columns. For example, in the next statement, you select the values of one column, income, divide each value by 12, and display the output in another column. The SELECT statement can use almost any type of expression. If more than one value is returned, the items are inserted into an array.

This SELECT statement uses the yearly income values divided by 12 to calculate the corresponding values for monthlysalary:

5 rows returned

This SELECT statement performs an addition operation that adds a bonus of 5000 to income to return salarywithbonus:

5 rows returned

#### Identifying tables and their columns

The FROM clause can contain one table only (that is, joins are not supported). The table is specified by its name, which may be followed by an optional alias. The table can be referenced in the other clauses either by its name or its alias. As we will see later, sometimes the use of the table name or alias is mandatory. However, for table



columns, the use of the table name or alias is optional. For example, here are three ways to write the same query:

```
sql-> SELECT Users.lastname, age FROM Users;
+-----+
| lastname | age |
+-----+
| Scully | 47 |
| Smith | 38 |
| Morgan | 38 |
| Anderson | 35 |
| Morrison | 25 |
+-----+
```

5 rows returned

To identify the table Users with the alias u:

```
sql-> SELECT lastname, u.age FROM Users u ;
```

The keyword AS can optionally be used before an alias. For example, to identify the table Users with the alias People:

```
sql-> SELECT People.lastname, People.age FROM Users AS People;
```

#### Filtering Results

You can filter query results by specifying a filter condition in the WHERE clause. Typically, a filter condition consists of one or more comparison expressions connected through logical operators AND or OR. The comparison operators are also supported: =, !=, >, >=, <, and <=.

This guery filters results to return only users whose first name is John:

2 rows returned

To return users whose calculated monthlysalary is greater than 6000:

```
sql-> SELECT id, lastname, income, income/12 AS monthlysalary
FROM Users WHERE income/12 > 6000;
+---+----+
| id | lastname | income | monthlysalary |
+---+----+
| 5 | Scully | 400000 | 33333 |
| 4 | Smith | 80000 | 6666 |
```



```
| 2 | Anderson | 100000 | 8333 |
| 1 | Morrison | 100000 | 8333 |
+---+
```

5 rows returned

To return users whose age is between 30 and 40 or whose income is greater than 100.000:

```
sql-> SELECT lastname, age, income FROM Users
WHERE age >= 30 and age <= 40 or income > 100000;
+-----+
| lastname | age | income |
+-----+
| Smith | 38 | 80000 |
| Morgan | 38 | NULL |
| Anderson | 35 | 100000 |
| Scully | 47 | 400000 |
+-----+
```

4 rows returned

You can use parenthesized expressions to alter the default precedence among operators. For example:

To return the users whose age is greater than 40 and either their age is less than 30 or their income is greater or equal than 100,000:

```
sql-> SELECT id, lastName FROM Users WHERE
(income >= 100000 or age < 30) and age > 40;
+---+---+
| id | lastName |
+---+----+
| 5 | Scully |
+---+----+
```

1 row returned

You can use the IS NULL condition to return results where a field column value is set to SQL NULL (SQL NULL is used when a non-JSON field is set to null):

```
sql-> SELECT id, lastname from Users WHERE income IS NULL;
+---+---+
| id | lastname |
+---+---+
| 3 | Morgan |
+---+-----+
```

1 row returned



You can use the IS NOT NULL condition to return column values that contain non-null data:

```
sql-> SELECT id, lastname from Users WHERE income IS NOT NULL;
+---+----+
| id | lastname |
+---+----+
| 4 | Smith |
| 1 | Morrison |
| 5 | Scully |
| 2 | Anderson |
+---+-----+
```

4 rows returned

#### **Grouping Results**

Use the GROUP BY clause to group the results by one or more table columns. Typically, a GROUP BY clause is used in conjunction with an aggregate expression such as COUNT, SUM, and AVG.



You can use the GROUP BY clause only if there exists an index that sorts the rows by the grouping columns.

For example, this query returns the average income of users, based on their age.

#### **Ordering Results**

Use the ORDER BY clause to order the results by a primary key column or a non-primary key column.



4 rows returned

You can use ORDER BY only if you are selecting by the table's primary key, or if there is an index that sorts the table's rows in the desired order.

To order by using a primary key column (id), specify the sort column in the ORDER BY clause:

To order by a non-primary key column, first create an index on the column of interest. For example, to use column lastname for ordering, create an index on that column, before using it in your ORDER BY clause:

```
sql-> CREATE INDEX idx1 on Users(lastname);
Statement completed successfully
sql-> SELECT id, lastname FROM Users ORDER BY lastname;
+---+---+
| id | lastname |
+---+---+
| 2 | Anderson |
| 3 | Morgan |
| 1 | Morrison |
| 5 | Scully |
| 4 | Smith |
+---+-----+
5 rows returned
```

Using this example data, you can order by more than one column if you create an index on the columns. (If our table had used more than one column for its primary key, then you can order by multiple columns using the primary keys.) For example, to order users by age and income.



```
5 rows returned
```

Creating a single index from two columns in the order you use them (age, income in this example), has some limits. The first column name (age) becomes the main sort item for the new index. You can use idx2 index to order by age only, but neither by income only, nor by income first and age second.

5 rows returned

To learn more about indexes see Working With Indexes.

By default, sorting is performed in ascending order. To sort in descending order use the DESC keyword in the ORDER BY clause:

```
sql-> SELECT id, lastname FROM Users ORDER BY id DESC;
+---+----+
| id | lastname |
+---+----+
| 5 | Scully |
| 4 | Smith |
| 3 | Morgan |
| 2 | Anderson |
| 1 | Morrison |
+---+-----+
```

5 rows returned

#### Limiting and Offsetting Results

Use the LIMIT clause to limit the number of results returned from a SELECT statement. For example, if there are 1000 rows in the Users table, limit the number of rows to return by specifying a LIMIT value. For example, this statement returns the first four ID rows from the table:



```
| 4 | Peter | Smith | 38 | 80000 | +---+
```

4 rows returned

To return only results 3 and 4 from the 10000 rows use the LIMIT clause to indicate 2 values, and the OFFSET clause to specify where the offset begins (after the first two rows). For example:

2 rows returned



We recommend using LIMIT and OFFSET with an ORDER BY clause. Otherwise, the results are returned in a random order, producing unpredictable results.

#### **Using External Variables**

Using external variables lets a query to written and compiled once, and then run multiple times with different values for the external variables. Binding the external variables to specific values is done through APIs, which you use before executing the query.

You must declare external variables in your SQL query before referencing them in the SELECT statement. For example:

```
DECLARE $age integer;
SELECT firstname, lastname, age
FROM Users
WHERE age > $age;
```

If the variable \$age is set to value 39, the result of the above query is:

+-		+-		+-		+
	firstname				_	
+-		+-		+-		+
	Dana		-			



## Working with complex data

In this chapter, we present query examples that use complex data types (arrays, maps, records). To follow along with the examples, get the  ${\tt Examples}$  download from here and run the  ${\tt SQLAdvancedExamples}$  script found in the  ${\tt sql}$  folder. This script creates the table and imports the data used.

#### SQLAdvancedExamples Script

The SQLAdvancedExamples script creates the following table:

```
CREATE TABLE Persons (
  id integer,
  firstname string,
  lastname string,
  age integer,
  income integer,
  lastLogin timestamp(4),
  address record(street string,
                 city string,
                  state string,
                  phones array (record (type enum (work, home),
                                      areacode integer,
                                      number integer
                ),
  connections array(integer),
  expenses map(integer),
  primary key (id)
);
```

The script also imports the following table rows:

```
},
 "connections":[2, 3],
  "expenses":{"food":1000, "gas":180}
}
  "id":2,
 "firstname": "John",
 "lastname": "Anderson",
 "age":35,
 "income":100000,
 "lastLogin": "2016-11-28T13:01:11.2088",
  "address":{"street":"187 Hill Street",
             "city": "Beloit",
             "state":"WI",
             "zipcode" : 53511,
             "phones":[{"type":"home", "areacode":339,
             "number":1684972}]
  "connections":[1, 3],
  "expenses": { "books": 100, "food": 1700, "travel": 2100 }
}
 "id":3,
 "firstname": "John",
 "lastname": "Morgan",
 "age":38,
 "income":100000000,
 "lastLogin": "2016-11-29T08:21:35.4971",
  "address":{"street":"187 Aspen Drive",
             "city": "Middleburg",
             "state":"FL",
             "phones":[{"type":"work", "areacode":305,
                         "number":1234079},
                       {"type":"home", "areacode":305,
                       "number":2066401}
 "connections":[1, 4, 2],
 "expenses":{"food":2000, "travel":700, "gas":10}
}
 "id":4,
 "firstname": "Peter",
 "lastname": "Smith",
 "age":38,
 "income":80000,
  "lastLogin": "2016-10-19T09:18:05.5555",
  "address":{"street":"364 Mulberry Street",
             "city": "Leominster",
             "state": "MA",
             "phones":[{"type":"work", "areacode":339,
                         "number":4120211},
```

```
{"type": "work", "areacode": 339,
                        "number":8694021},
                       {"type": "home", "areacode": 339,
                        "number":1205678},
                       {"type": "home", "areacode": 305,
                        "number":8064321}
                      1
            },
  "connections":[3, 5, 1, 2],
  "expenses":{"food":6000, "books":240, "clothes":2000, "shoes":1200}
}
 "id":5,
 "firstname": "Dana",
 "lastname": "Scully",
 "age":47,
 "income":400000,
 "lastLogin": "2016-11-08T09:16:46.3929",
  "address":{"street":"427 Linden Avenue",
             "city": "Monroe Township",
             "state":"NJ",
             "phones":[{"type":"work", "areacode":201,
                        "number":3213267},
                       {"type": "work", "areacode": 201,
                        "number":8765421},
                       {"type": "home", "areacode": 339,
                        "number":3414578}
                      1
           },
  "connections": [2, 4, 1, 3],
  "expenses":{"food":900, "shoes":1000, "clothes":1500}
}
```

#### You run the SQLAdvancedExamples script using the load command:

```
> cd <installdir>/examples/sql
> java -jar <KVHOME>/lib/sql.jar -helper-hosts <host>:<port> \
-store <storename> load \
-file <KVHOME>/examples/sql/SQLAdvancedExamples.cli
```



#### Note:

The Persons table schema models people that can be connected to other people in the table. All connections are stored in the "connections" column, which consists of an array of integers. Each integer is an ID of a person with whom the subject is connected. The entries in the "connections" array are sorted in descending order, indicating the strength of the connection. For example, looking at the record for person 3, we see that John Morgan has these connections: [1, 4, 2]. The order of the array elements specifies that John is most strongly connected with person 1, less connected with person 4, and least connected with person 2.

Records in the Persons table also include an "expenses" column, declared as an integer map. For each person, the map stores key-value pairs of string item types and integers representing money spent on the item. For example, one record has these expenses: {"food":900, "shoes":1000, "clothes":1500}, other records have different items. One benefit of modelling expenses as a map type is to facilitate the categories being different for each person. Later, we may want to add or delete categories dynamically, without changing the table schema, which maps readily support. An item to note about this map is that it is an integer map always contains key-value pairs, and keys are always strings.

#### Working with Timestamps

To specify a timestamp value in a query, provide it as a string, and cast it to a Timestamp data type. For example:

```
sql-> SELECT id, firstname, lastname FROM Persons WHERE
lastLogin = CAST("2016-10-19T09:18:05.5555" AS TIMESTAMP);
+---+----+
| id | firstname | lastname |
+---+----+
| 4 | Peter | Smith |
+---+------+
```

1 row returned

Timestamp queries often involve a range of time, which requires multiple casts:



3 rows returned

You can also use various Timestamp functions to return specific time and date values from the Timestamp data. For example:

id	+   firstname +	lastname	Year	Month	Day	Hour	Minute
		Morgan					21
2	John	Anderson	2016	11	28	13	1
4	Peter	Smith	2016	10	19	9	18
5	Dana	Scully	2016	11	8	9	16
1	David	Morrison	2016	10	29	18	43
	+	+	+	+	·	+	++

#### Alternatively, use the EXTRACT function:

```
sql-> SELECT id, firstname, lastname,
    EXTRACT(YEAR FROM lastLogin) AS Year,
    EXTRACT(MONTH FROM lastLogin) AS Month,
    EXTRACT(DAY FROM lastLogin) AS Day,
    EXTRACT(HOUR FROM lastLogin) AS Hour,
    EXTRACT(MINUTE FROM lastLogin) AS Minute
FROM Persons;
```

++	+	+	+	<b></b>		++
id   firstname	lastname	Year	Month	Day	Hour	Minute
3   John	Morgan	2016	11	29	8	21
	Smith	'			9	- 1
	Morrison   Anderson	'		29     28	-	43   1
	Scully	'		20		
++	+	+	+	<del></del> +		++

5 rows returned
sql->

#### **Working With Arrays**

You can use slice or filter steps to select elements out of an array. We start with some examples using slice steps.



To select and display the second connection of each person, we use this query:

```
sql-> SELECT lastname, connections[1]
AS connection FROM Persons;
+-----+
| lastname | connection |
+----+
| Scully | 2 |
| Smith | 4 |
| Morgan | 2 |
| Anderson | 2 |
| Morrison | 2 |
+-----+
```

5 rows returned

In the example, the slice step [1] is applied to the connections array. Since array elements start with 0, 1 selects the second connection value.

You can also use a slice step to select all array elements whose positions are within a range: [low:high], where low and high are expressions to specify the range boundaries. You can omit low and high expressions if you do not require a low or high boundary.

For example, the following query returns the lastname and the first 3 connections of person 5 as strongconnections:

1 row returned

In the above query for Person 5, the path expression <code>connections[0:2]</code> returns the person's first 3 connections. Here, the range is [0:2], so 0 is the low expression and 2 is the high. The path expression returns its result as a list of 3 items. The list is converted to an array (a single item) by enclosing the path expression in an array-constructor expression ([]). The array constructor creates a new array containing the three connections. Notice that although the query shell displays the elements of this constructed array vertically, the number of rows returned by this query is 1.

Use of the array constructor in the select clause is optional. If no array constructor is used, an array will still be constructed, but only if the select-clause expression does indeed return more than one item. If exactly one item is returned, the result will contain just that one item. If the expression returns nothing (an empty result), NULL is used as the result. This behavior is illustrated in the next example, which we will run with and without an array constructor.

As mentioned above, you can omit the low or high expression when specifying the range for a slice step. For example the following query specifies a range of [3:] which



returns all connections after the third one. Notice that for persons having only 3 connections or less, an empty array is constructed and returned due to the use of the array constructor.

To fully illustrate this behavior, we display this output in mode JSON because the COLUMN mode does not differentiate between a single item and an array containing a single item.

```
sql-> mode JSON
Query output mode is JSON
sql-> SELECT id, [connections[3:]] AS weakConnections FROM Persons;
{"id":3,"weakConnections":[]}
{"id":4,"weakConnections":[2]}
{"id":2,"weakConnections":[]}
{"id":5,"weakConnections":[3]}
{"id":1,"weakConnections":[]}
```

Now we run the same query, but without the array constructor. Notice how single items are not contained in an array, and for rows with no match, NULL is returned instead of an empty array.

```
sql-> SELECT id, connections[3:] AS weakConnections FROM Persons;
{"id":2,"weakConnections":null}
{"id":3,"weakConnections":null}
{"id":4,"weakConnections":2}
{"id":5,"weakConnections":3}
{"id":1,"weakConnections":null}

5 rows returned
sql-> mode COLUMN
Query output mode is COLUMN
sql->
```

As a last example of slice steps, the following query returns the last 3 connections of each person. In this query, the slice step is [size(\$)-3:]. In this expression, the \$ is an implicitly declared variable that references the array that the slice step is applied to. In this example, \$ references the connections array. The size() built-in function returns the size (number of elements) of the input array. So, in this example, size(\$) is the size of the current connections array. Finally, size(\$)-3 computes the third position from the end of the current connections array.



			4			
	1		2	 	 	

5 rows returned

We now turn our attention to filter steps on arrays. Like slice steps, filter steps also use the square brackets ([]) syntax. However, what goes inside the [] is different. With filter steps there is either nothing inside the [] or a single expression that acts as a condition (returns a boolean result). In the former case, all the elements of the array are selected (the array is "unnested"). In the latter case, the condition is applied to each element in turn, and if the result is true, the element is selected, otherwise it is skipped. For example:

The following query returns the id and connections of persons who are connected to person 4:

2 rows returned

In the above query, the expression p.connections[] returns all the connections of a person. Then, the =any operator returns true if this sequence of connections contains the number 4.

The following query returns the id and connections of persons who are connected with any person having an id greater than 4:

```
sql-> SELECT id, connections FROM Persons p
WHERE p.connections[] >any 4;
+---+---+
| id | connections |
+---+----+
| 4 | 3 |
```



		5		
		1		
		2		
+	+-		 	+

1 row returned

5 rows returned

The following query returns, for each person, the person's last name and the phone numbers with area code 339:

```
sql-> SELECT lastname,
[ p.address.phones[$element.areacode = 339].number ]
AS phoneNumbers FROM Persons p;
+----+
| lastname | phoneNumbers |
+----+
| Scully | 3414578
+----+
| Smith | 4120211
     8694021
    | 1205678
+----+
| Morgan |
+----+
| Anderson | 1684972
+----+
| Morrison |
+----+
```

In the above query, the filter step [\$element.areacode = 339] is applied to the phones array of each person. The filter step evaluates the condition \$element.areacode = 339 on each element of the array. This condition expression uses the implicitly declared variable \$element, which references the current element of the array. An empty array is returned for persons that do not have any phone number in the 339 area code. If we wanted to filter out such persons from the result, we would write the following query:



```
3 rows returned
```

The previous query contains the path expression p.address.phones.areacode. In that expression, the field step .areacode is applied to an array field (phones). In this case, the field step is applied to each element of the array in turn. In fact, the path expression is equivalent to p.address.phones[].areacode.

In addition to the implicitly-declared \$ and \$element variables, the condition inside a filter step can also use the \$pos variable (also implicitly declared). \$pos references the position within the array of the current element (the element on which the condition is applied). For example, the following query selects the "interesting" connections of each person, where a connection is considered interesting if it is among the 3 strongest connections and connects to a person with an id greater or equal to 4.

5 rows returned

Finally, two arrays can be compared with each other using the usual comparison operators (=, !=, >, >=, >, and >=). For example the following query constructs the array [1,3] and selects persons whose connections array is equal to [1,3].

```
sql-> SELECT lastname FROM Persons p
WHERE p.connections = [1,3];
+-----+
| lastname |
+-----+
| Anderson |
+-----+
1 row returned
```



#### Working with Records

1 row returned

You can use a field step to select the value of a field from a record. For example, to return the id, last name, and city of persons who reside in Florida:

In the above query, the path expression p.address.state consists of 2 field steps: .address selects the address field of the current row (rows can be viewed as records, whose fields are the row columns), and .state selects the state field of the current address.

The example record contains an array of phone numbers. You can form queries against that array using a combination of path steps and sequence comparison operators. For example, to return the last name of persons who have a phone number with area code 423:

```
sql-> SELECT lastname FROM Persons
p WHERE p.address.phones.areacode =any 423;
+-----+
| lastname |
+-----+
| Morrison |
+-----+
1 row returned
```

In the above query, the path expression p.address.phones.areacode returns all the area codes of a person. Then, the =any operator returns true if this sequence of area codes contains the number 423. Notice also that the field step .areacode is applied to an array field (phones). This is allowed if the array contains records or maps. In this case, the field step is applied to each element of the array in turn.

The following example returns all the persons who had three connections. Notice the use of [] after connections: it is an array filter step, which returns all the elements of the connections array as a sequence (it is unnesting the array).



	 +		3
4	Peter     	Smith   	3
5       	Dana       	Scully     	2

4 rows returned

This query can use ORDER BY to sort the results because the sort is being performed on the table's primary key. The next section shows sorting on non-primary key fields through the use of indexes.

For more examples of querying against data contained in arrays, see Working With Arrays.

#### Using ORDER BY to Sort Results

To sort the results from a SELECT statement using a field that is not the table's primary key, you must first create an index for the column of choice. For example, for the next table, to query based on a Timestamp and sort the results in descending order by the timestamp, create an index:

```
+---+----+
| id | firstname | lastLogin
+---+
| 3 | John | Morgan | 2016-11-29T08:21:35.4971 | 4 | Peter | Smith | 2016-10-19T09:18:05.5555 |
| 2 | John
          | Anderson | 2016-11-28T13:01:11.2088 |
| 5 | Dana
          | Scully | 2016-11-08T09:16:46.3929 |
| 1 | David | Morrison | 2016-10-29T18:43:59.8319 |
+---+
5 rows returned
sql-> CREATE INDEX tsidx1 on Persons (lastLogin);
Statement completed successfully
sgl-> SELECT id, firstname, lastname, lastLogin
FROM Persons ORDER BY lastLogin DESC;
+---+
| id | firstname | lastLogin
+---+
| 3 | John | Morgan | 2016-11-29T08:21:35.4971 |
          | Anderson | 2016-11-28T13:01:11.2088 |
| 2 | John
```

sql-> SELECT id, firstname, lastname, lastLogin FROM Persons;



+---+

5 rows returned

SQL for Oracle NoSQL Database can also sort query results by the values of nested records. To do so, create an index of the nested field (or fields). For example, you can create an index of address.state from the Persons table, and then order by state:

5 rows returned

To learn more about indexes, see Working With Indexes.

#### Working With Maps

The path steps applicable to maps are field and filter steps. Slice steps do not make sense for maps, because maps are unordered, and as a result, their entries do not have any fixed positions.

You can use a field step to select the value of a field from a map. For example, to return the lastname and the food expenses of all persons:

```
sql-> SELECT lastname, p.expenses.food
FROM Persons p;
+-----+
| lastname | food |
+-----+
| Morgan | 2000 |
| Morrison | 1000 |
| Scully | 900 |
| Smith | 6000 |
| Anderson | 1700 |
+-----+
```

5 rows returned

In the above query, the path expression p.expenses.food consists of 2 field steps: .expenses selects the expenses field of the current row and .food selects the value of the food field/entry from the current expenses map.



To return the lastname and amount spent on travel for each person who spent less than \$3000 on food:

```
sql-> SELECT lastname, p.expenses.travel
FROM Persons p WHERE p.expenses.food < 3000;
+-----+
| lastname | travel |
+-----+
| Scully | NULL |
| Morgan | 700 |
| Anderson | 2100 |
| Morrison | NULL |
+-----+</pre>
```

4 rows returned

Notice that NULL is returned for persons who did not have any travel expenses.

Filter steps are performed using either the <code>.values()</code> or <code>.keys()</code> path steps. To select values of map entries, use <code>.values(<cond>)</code>. To select keys of map entries, use <code>.keys(<cond>)</code>. If no condition is used in these steps, all the values or keys of the input map are selected. If the steps do contain a condition expression, the condition is evaluated for each entry, and the value or key of the entry is selected/skipped if the result is true/false.

The implicitly-declared variables \$key and \$value can be used inside a map filter condition. \$key references the key of the current entry and \$value references the associated value. Notice that, contrary to arrays, the \$pos variable can not be be used inside map filters (because map entries do not have fixed positions).

To show, for each user, their id and the expense categories where they spent more than \$1000:

```
sql-> SELECT id, p.expenses.keys($value > 1000) as Expenses
from Persons p;
+---+
| id | Expenses
+---+
| 4 | clothes |
| | food
  | shoes
| 3 | food
+---+
| 2 | food
| | travel
+---+
| 5 | clothes
+---+
| 1 | NULL |
+---+
```

To return the id and the expense categories in which the user spent more than they spent on clothes, use the following filter step expression. In this query, the context-item



variable (\$) appearing in the filter step expression [\$value > \$.clothes] refers to the expenses map as a whole.

To return the id and expenses data of any person who spent more on any category than what they spent on food:

2 rows returned

To return the id of all persons who consumed more than \$2000 in any category other than food:

```
sql-> SELECT id FROM Persons p
WHERE p.expenses.values($key != "food") >any 2000;
+----+
| id |
+----+
| 2 |
+----+
```

1 row returned

#### Using the size() Function

The size function can be used to return the size (number of fields/entries) of a complex item (record, array, or map). For example:

To return the id and the number of phones that each person has:

5 rows returned

To return the id and the number of expenses categories for each person: has:

5 rows returned

To return for each person their id and the number of expenses categories for which the expenses were more than 2000:

```
sql-> SELECT id, size([p.expenses.values($value > 2000)]) AS
expensiveCategories FROM Persons p;
+---+---+
| id | expensiveCategories |
+---+---+
| 3 | 0 |
| 2 | 1 |
| 5 | 0 |
| 1 | 0 |
| 4 | 1 |
```



5 rows returned



4

### Working with JSON

This chapter provides examples on working with JSON data. If you want to follow along with the examples, get the <code>Examples</code> download from here and run the <code>SQLJSONExamples</code> script found in the <code>sqlfolder</code>. This creates the table and imports the data used.

JSON data is written to JSON data columns by providing a JSON object. This object can contain any valid JSON data. The input data is parsed and stored internally as Oracle NoSQL Database datatypes:

- When numbers are encountered, they are converted to integer, long, or double items, depending on the actual value of the number (float items are not used for JSON).
- Strings in the input text are mapped to string items.
- Boolean values are mapped to boolean items.
- JSON nulls are mapped to JSON null items.
- When an array is encountered in the input text, an array item is created whose type is
   Array (JSON). This is done unconditionally, no matter what the actual contents of the
   array might be.
- When a JSON object is encountered in the input text, a map item is created whose type is Map (JSON), unconditionally.



There is no JSON equivalent to the TIMESTAMP datatype, so if input text contains a string in the TIMESTAMP format it is simply stored as a string item in the JSON column.

The remainder of this chapter provides an overview to querying JSON data.

### **SQLJSONExamples Script**

The SQLJSONExample is available to illustrate JSON usage. This script creates the following table:

```
create table if not exists JSONPersons (
  id integer,
  person JSON,
  primary key (id)
);
```

The script imports the following table rows. Notice that the content for the person column, which is of type JSON contains a JSON object. That object contains a series of fields which

represent our person. We have deliberately included inconsistent information in this example so as to illustrate how to handle various queries when working with JSON data.

```
"id":1,
  "person" : {
      "firstname": "David",
      "lastname": "Morrison",
      "age":25,
      "income":100000,
      "lastLogin": "2016-10-29T18:43:59.8319",
      "address":{"street":"150 Route 2",
                 "city": "Antioch",
                 "state": "TN",
                 "zipcode" : 37013,
                 "phones":[{"type":"home", "areacode":423,
                             "number":8634379}]
                },
      "connections": [2, 3],
      "expenses":{"food":1000, "gas":180}
}
  "id":2,
  "person" : {
      "firstname": "John",
      "lastname": "Anderson",
      "age":35,
      "income":100000,
      "lastLogin": "2016-11-28T13:01:11.2088",
      "address":{"street":"187 Hill Street",
                 "city": "Beloit",
                 "state":"WI",
                 "zipcode" : 53511,
                 "phones":[{"type":"home", "areacode":339,
                             "number":1684972}]
                },
      "connections":[1, 3],
      "expenses": { "books": 100, "food": 1700, "travel": 2100}
}
  "id":3,
  "person" : {
      "firstname":"John",
      "lastname": "Morgan",
      "age":38,
      "income":100000000,
      "lastLogin" : "2016-11-29T08:21:35.4971",
      "address":{"street":"187 Aspen Drive",
                 "city": "Middleburg",
                 "state":"FL",
```

```
"phones":[{"type":"work", "areacode":305,
                             "number":1234079},
                            {"type": "home", "areacode": 305,
                             "number":2066401}
                },
      "connections":[1, 4, 2],
      "expenses":{"food":2000, "travel":700, "gas":10}
  }
}
 "id":4,
  "person": {
      "firstname": "Peter",
      "lastname": "Smith",
      "age":38,
      "income":80000,
      "lastLogin" : "2016-10-19T09:18:05.5555",
      "address":{"street":"364 Mulberry Street",
                 "city": "Leominster",
                 "state": "MA",
                 "phones":[{"type":"work", "areacode":339,
                             "number":4120211},
                           {"type":"work", "areacode":339,
                            "number":8694021},
                           {"type": "home", "areacode": 339,
                            "number":1205678},
                           {"type": "home", "areacode": 305,
                            "number":8064321}
                },
      "connections":[3, 5, 1, 2],
      "expenses":{"food":6000, "books":240, "clothes":2000,
                  "shoes":1200}
}
  "id":5,
  "person" : {
      "firstname": "Dana",
      "lastname": "Scully",
      "age":47,
      "income":400000,
      "lastLogin" : "2016-11-08T09:16:46.3929",
      "address":{"street":"427 Linden Avenue",
                 "city": "Monroe Township",
                 "state": "NJ",
                 "phones":[{"type":"work", "areacode":201,
                             "number":3213267},
                           {"type": "work", "areacode": 201,
                            "number":8765421},
                           {"type": "home", "areacode": 339,
                            "number":3414578}
```

You run the SQLJSONExamples script using the load command:

```
> cd <installdir>/examples/sql
> java -jar <KVHOME>/lib/sql.jar -helper-hosts <host>:<port> \
-store <storename> load \
-file <KVHOME>/examples/sql/SQLJSONExamples.cli
```

### **Basic Queries**

Because JSON is parsed and stored internally in native data formats with Oracle NoSQL Database, querying JSON data is no different than querying data in other column types. See Simple SELECT Queries and Working with complex data for introductory examples of how to form these queries.

In our JSONPersons example, all of the data for each person is contained in a column of type JSON called person. This data is presented as a JSON object, and mapped internally into a Map(JSON) type. You can query information in this column as you would query a Map of any other type. For example:



1   Morrison	25
	38
6   NULL	NULL
7   NULL	NULL
TT	-+

<sup>7</sup> rows returned

The last two rows in returned from this query contain all NULLs. This is because those rows were populated using JSON objects that are different than the objects used to populate the rest of the table. This capability of JSON is both a strength and a weakness. As a plus, you can modify your schema easily. However, if you are not careful, you can end up with tables containing dissimilar data in both large and small ways.

Because the JSON object is stored as a map, you can use normal map step functions on the column. For example:

```
sql-> SELECT id, j.person.expenses.keys($value > 1000) as Expenses
from JSONPersons j;
| id | Expenses |
+---+
| 3 | food
+---+
| 2 | food
 | travel
+---+
| 4 | clothes |
  | food
  | shoes
+---+
| 6 | NULL
+---+
| 5 | clothes
+---+
| 7 | NULL
+---+
| 1 | NULL
+---+
```

Here, id 1 is NULL because that user had no expenses greater than \$1000, while id 6 and 7 are NULL because they have no j.person.expenses field.

## Using WHERE EXISTS with JSON

As we saw in the previous section, different rows in the same table can have dissimilar information in them when a column type is JSON. To identify whether desired information exists for a given JSON column, use the EXISTS operator.



<sup>7</sup> rows returned

For example, some of the JSON persons have a zip code entered for their address, and others do not. Use this query to see all the users with a zipcode:

2 rows returned

When querying data for inconsistencies, it is often more useful to see all rows where information is missing by using WHERE NOT EXISTS:

1 row returned



## Seeking NULLS in Arrays

All arrays found in a JSON input stream are stored internally as ARRAY(JSON). This means that it is possible for the array to have inconsistent types for its members.

In our example, the phones array for user id 4 contains a null element:

```
sql-> SELECT j.person.address.phones FROM JSONPersons j WHERE j.id=4;
+----+
     phones
+----+
| areacode | 339
| number | 4120211 |
| areacode | 339
| number | 8694021 |
| type | work |
| areacode | 339
| number | 1205678 |
| type | home
| null
| areacode | 305
| number | 8064321 |
| type | home
```

A way to discover this in your table is to examine the phones array for null values:

```
sql-> SELECT id, j.person.address.phones FROM JSONPersons j
WHERE j.person.address.phones[] =any null;
+---+
| id | phones
  4 | areacode | 339
    | number | 4120211 |
    | areacode | 339
    | number | 8694021 |
    | areacode | 339
    | number | 1205678 |
    | type | home
    | null
    | areacode | 305
    | number | 8064321 |
    | type | home |
```



```
1 row returned
```

Notice the use of the array filter step ([]) in the previous query. This is needed to unpack the array into a sequence so that the =any comparison operator can be used with it.

## **Examining Data Types JSON Columns**

The example data contains a couple of rows with unusual data:

```
{
    "id":6,
    "person" : {
        "mynumber":5,
        "myarray":[1,2,3,4]
    }
}

{
    "id":7,
    "person" : {
        "mynumber":"5",
        "myarray":["1","2","3","4"]
    }
}
```

You can locate them using the query:

2 rows returned



However, notice that these two rows actually contain numbers stored as different types. ID 6 stores integers while ID 7 stores strings. You can select a row based on its type:

Notice that if you use IS NOT OF TYPE then every row in the table is returned except id 6. This is because for all the other rows, j.person.mynumber evaluates to jnull, which is not an integer.

```
sql-> SELECT id FROM JSONPersons j
WHERE j.person.mynumber IS NOT OF TYPE (integer);
+---+
| id |
+---+
| 3 |
| 2 |
| 5 |
| 4 |
| 1 |
| 7 |
+---+
```

6 rows returned

1 row returned

To solve this problem, also check for the existence of j.person.mynumber:

```
sql-> SELECT id from JSONPersons j WHERE EXISTS j.person.mynumber
and j.person.mynumber IS NOT OF TYPE (integer);
+----+
| id |
+----+
| 7 |
+----+
```



You can also perform type checking based on the type of data contained in the array. Recall that our rows contain arrays with integers and arrays with strings. You can return the row with just the array of strings using:

1 row returned

6 rows returned

Here, we use the array filter step ([]) in the WHERE clause to unpack the array into a sequence. This allows is-of-type to iterate over the sequence, checking the type of each element. If every element in the sequence matches the identified type (string, in this case), then the is-of-type returns true.

Also notice that the query uses the + cardinality modifier. This means that is-of-type will return true only if the input sequence (myarray[], in this case) contains ONE OR MORE elements that match the identified type (string). If we used \*, then 0 or more elements would have to match the identified type in order for true to return. Because our table contains a mix of rows with different schema, the result is that every row except id 6 is returned:

```
sql-> SELECT id, j.person.myarray FROM JSONPersons j
WHERE j.person.myarray[] IS OF TYPE (string*);
+---+
| id | myarray |
+---+
| 3 | NULL
+---+
| 5 | NULL
+---+
| 1 | NULL
+---+
| 7 | 1
| | 2
  | 3
  | 4
| 4 | NULL
| 2 | NULL
+---+
```



Finally, if we do not provide a cardinality modifier at all, then is-of-type returns true if ONE AND ONLY one member of the input sequence matches the identified type. In this example, the result is that no rows are returned.

```
sql-> SELECT id, j.person.myarray FROM JSONPersons j
WHERE j.person.myarray[] IS OF TYPE (string);
0 row returned
```

## Using Map Steps with JSON Data

On import, Oracle NoSQL Database stores JSON objects as MAP(JSON). This means you can use map filter steps with your JSON objects.

For example, if you want to visually examine the JSON fields in use by your rows:

```
sql-> SELECT id, j.person.keys() FROM JSONPersons j;
+---+
| id | Column 2
+---+
| 4 | address
| | age
  | connections
   | expenses
   | firstname
   | income
   | lastLogin
   | lastname
+---+
| 6 | myarray
| | mynumber
+---+
| 3 | address
 | age
   | connections
| expenses
   | firstname
   | income
   | lastLogin
   | lastname
 5 | address
   age
   | connections
   expenses
   | firstname
   | income
   | lastLogin
   | lastname
| 1 | address
   | age
   | connections
   expenses
```



7 rows returned

## **Casting Datatypes**

You can cast one data type to another using the cast expression.

In JSON, casting is particularly useful for timestamp information because JSON has no equivalent to the Oracle NoSQL Database Timestamp data type. Instead, the timestamp information is carried in a JSON object as a string. To work with it as a Timestamp, use <code>cast</code>.

In Working with Timestamps we showed how to work with the timestamp data type. In this case, what you do is no different except you must cast both sides of the expression. Also, because the left side of the expression is a sequence, you must specify a type quantifier (\* in this case):

3 rows returned



As another example, you can cast to an integer and then operate on that number:

If you want to operate on just the row that contains the number as a string, use IS OF TYPE:

## **Using Searched Case**

A searched case expression can be helpful in identifying specific problems with the JSON data in your JSON columns. The example data we have been using in this chapter sometimes provides a JSONPersons.address field, and sometimes it does not. When an address is present, sometimes it provides a zipcode, and sometimes it does not. We can use a searched case expression to identify and describe the specific problem with each row.

```
sql-> SELECT id,
CASE
  WHEN NOT EXISTS j.person.address
  THEN j.person.keys()
  WHEN NOT EXISTS j.person.address.zipcode
  THEN "No Zipcode"
  ELSE j.person.address.zipcode
END
FROM JSONPersons j;
+---+
| id | Column_2 |
+---+
| 4 | No Zipcode
+---+
| 3 | No Zipcode
+---+
| 5 | No Zipcode
+---+
| 1 | 37013
+---+
```



1	7	myarray   mynumber
+-	6	myarray   mynumber
	2	53511

<sup>7</sup> rows returned

We can improve the report by adding a third column that uses a second searched case expression:

```
sql-> SELECT id,
CASE
  WHEN NOT EXISTS j.person.address
  THEN "No Address"
  WHEN NOT EXISTS j.person.address.zipcode
  THEN "No Zipcode"
  ELSE j.person.address.zipcode
END,
CASE
  WHEN NOT EXISTS j.person.address
  THEN j.person.keys()
  ELSE j.person.address
END
FROM JSONPersons j;
| id | Column_2 | Column_3 |
| 3 | No Zipcode
                  | city | Middleburg
                  | phones
   | areacode | 305
                      number | 1234079
                      type | work
                      areacode | 305
                      number | 2066401
                     type | home
                   street
                          | Beloit
2 | 53511
                   | city
                   | phones
                   | areacode | 339
                      number | 1684972
                   | type | home
                   | state | WI
                           | 187 Hill Street
                   street
                  --+-----
             | city
| 5 | No Zipcode
                           | Monroe Township
```

| phones

		areacode   201
+	37013	city
7 	No Address	myarray     mynumber
	No Zipcode	city
6     6	No Address	myarray     mynumber

7 rows returned

Finally, it is possible to nest search case expressions. Our sample data also has a spurious null in the phones array (see id 4). We can report that in the following way (output is modified slightly to fit in the space allowed):

```
sql-> SELECT id,
CASE
  WHEN EXISTS j.person.address
  THEN
    CASE
      WHEN EXISTS j.person.address.zipcode
      THEN
         CASE
           WHEN j.person.address.phones[] =any null
           THEN "Zipcode exists but null in the phones array"
           ELSE j.person.address.zipcode
      WHEN j.person.address.phones[] =any null
      THEN "No zipcode and null in phones array"
      ELSE "No zipcode"
  ELSE "No Address"
END,
CASE
  WHEN NOT EXISTS j.person.address
  THEN j.person.keys()
  ELSE j.person.address
END
FROM JSONPersons j;
| id | Column 2 | Column 3
+---+
                    | city | Middleburg
| 3 | No zipcode
                     phones
                     | areacode | 305
                     | number | 1234079
                    | type | work
                     | areacode | 305
                    | number | 2066401
                    | type | home
                    | city | Beloit
| 2 | 53511
                     | phones
                    | areacode | 339
                        number | 1684972
                    type | home
                    | city | Monroe Township
| phones
| 5 | No zipcode
                  | areacode | 201
```

	 	number     type	
	 	areacode     number     type	8765421
	 		3414578
1	37013 		423 8634379 home TN 150 Route 2
7	No Address	myarray   mynumber	
4		type	339 4120211 work 339 8694021 work 339 1205678 home null
		street	304 Mulberry Street

7 rows returned

5

## Working With GeoJSON Data

The GeoJSON specification (https://tools.ietf.org/html/rfc7946) defines the structure and content of JSON objects representing geographical shapes on earth (called geometries). Oracle NoSQL Database implements several functions that interpret JSON geometry objects. The functions also let you search table rows containing geometries that satisfy certain conditions. Search is made efficient through the use of special indexes, as described in the *SQL Reference Guide*.



Support for GeoJson data is available only in the Oracle NoSQL Database Enterprise Edition.

### **Geodetic Coordinates**

As described, all kinds of geometries are specified in terms of a set of positions. However, for line strings and polygons, the actual geometrical shape is formed by lines connecting their positions. The GeoJSON specification defines a line between two points as the straight line that connects the points in the (flat) cartesian coordinate system, whose horizontal and vertical axes are the longitude and latitude, respectively. More precisely, the coordinates of every point on a line that does not cross the antimeridian between a point P1 = (lon1, lat1) and P2 = (lon2, lat2) can be calculated as:

```
P = (lon, lat) = (lon1 + (lon2 - lon1) * t, lat1 + (lat2 - lat1) * t)
```

with t being a real number, greater than or equal to 0, and less than or equal to 1.

Unlike the GeoJSON specification, the Oracle NoSQL Database uses a *geodetic* coordinate system, as defined in the World Geodetic System, WGS84, (https://gisgeography.com/wgs84-world-geodetic-system). A geodetic line between two points is the shortest line that can be drawn between the two points on the ellipsoidal surface of the earth.



### **GeoJSON Data Definitions**

The GeoJSON specification (https://tools.ietf.org/html/rfc7946) states that for a JSON object to be a geometry, it requires two fields, type and coordinates. The value of the type field specifies the kind of geometric shape the object describes. The value of the type field must be one of the following strings, corresponding to different kinds of geometries:

- Point
- LineSegment
- Polygon
- MultiPoint
- MultiLineString
- MultiPolygon
- GeometryCollection

The coordinates value is an array with elements that define the geometrical shape. An exception to this is the *GeometryCollection* type, which is described below. The coordinates value depends on the geometric shape, but in all cases, specifies a number of positions. A *position* defines a position on the surface of the earth as an array of two double numbers, where the first number is the longitude and the second number is the latitude. Longitude and latitude are specified as degrees and must range between -180 - +180 and -90 - +90, respectively.



The GeoJSON specification allows a third coordinate for the altitude of the position, but Oracle NoSQL Database does not support altitudes.



The kinds of geometries are defined as follows, each with an example of such an object:

**Point** — For type Point, the coordinates field is a single position:

```
{ "type" : "point", "coordinates" : [ 23.549, 35.2908 ] }
```

**LineString** — A LineString is one or more connected lines, with the end-point of one line being the start-point of the next. The coordinates field is an array of two or more positions. The first position is the start point of the first line, and each subsequent position is the end point of the previous line and the start of the next line. Lines can cross each other.

```
{
"type" : "LineString",
"coordinates" : [
[-121.9447, 37.2975],
[-121.9500, 37.3171],
[-121.9892, 37.3182],
[-122.1554, 37.3882],
[-122.2899, 37.4589],
[-122.4273, 37.6032],
[-122.4304, 37.6267],
[-122.3975, 37.6144]
]
}
```

**Polygon** — A polygon defines a surface area by specifying its outer perimeter and the perimeters of any potential holes inside the area. More precisely, a polygon consists of one or more linear rings, where (a) a linear ring is a closed LineString with four or more positions, (b) the first and last positions are equivalent, and they must contain identical values, (c) a linear ring is the boundary of a surface or the boundary of a hole in a surface, and (d) a linear ring must follow the right-hand rule with respect to the area it bounds. That is, positions for exterior rings must be ordered counterclockwise, and positions for holes must be ordered clockwise. Then, the coordinates field of a polygon must be an array of linear ring coordinate arrays, where the first must be the exterior ring, and any others must be interior rings.

The exterior ring bounds the surface, and the interior rings (if present) bound holes within the surface. The example below shows a polygon with no holes.

```
{
"type": "polygon",
"coordinates": [[
[23.48, 35.16],
[24.30, 35.16],
[24.30, 35.50],
[24.16, 35.61],
[23.74, 35.70],
[23.56, 35.60],
[23.48, 35.16]
]
]
]
}
```



**MultiPoint** — For type MultiPoint, the coordinates field is an array of two or more positions:

```
{
"type" : "MultiPoint",
"coordinates" : [
[-121.9447, 37.2975],
[-121.9500, 37.3171],
[-122.3975, 37.6144]
]
}
```

**MultiLineString** — For type MultiLineString, the coordinates member is an array of LineString coordinate arrays.

```
{
"type": "MultiLineString",
"coordinates": [
  [100.0, 0.0], [01.0, 1.0] ],
  [102.0, 2.0], [103.0, 3.0] ]
}
```

**MultiPolygon** — For type MultiPolygon, the coordinates member is an array of Polygon coordinate arrays.

```
"type": "MultiPolygon",
"coordinates": [
[
[102.0, 2.0],
[103.0, 2.0],
[103.0, 3.0],
[102.0, 3.0],
[102.0, 2.0]
]
],
ſ
[100.0, 0.0],
[101.0, 0.0],
[101.0, 1.0],
[100.0, 1.0],
[100.0, 0.0]
1
]
]
}
```

**GeometryCollection** — Instead of a coordinates field, a GeometryCollection has a geometries" field. The value of geometries is an array. Each element of this array is a

GeoJSON object whose kind is one of the six kinds defined above. In general, a GeometryCollection is a heterogeneous composition of smaller geometries.

```
{ "type": "GeometryCollection",
"geometries": [
{
"type": "Point",
"coordinates": [100.0, 0.0]
},
{"type": "LineString",
"coordinates": [ [101.0, 0.0], [102.0, 1.0] ]
}
]
}
```

#### Note:

The GeoJSON specification defines two additional kinds of entities, *Feature* and *FeatureCollection*. The Oracle NoSQL Database does not support these entities.

## Searching GeoJSON Data

The Oracle NoSQL Database has the following functions to use for searching GeoJSON data that has some relationship with a search geometry.

- boolean geo intersect(any\*, any\*)
- boolean geo inside(any\*, any\*)
- boolean geo\_within\_distance(any\*, any\*, double)
- boolean geo near(any\*, any\*, double)

In addition to the search functions, two other functions are available, and listed as the last two rows of the table:

Function	Туре	Details
<pre>geo_intersect(any*, any*)</pre>	boolea n	Raises an error at compile time if the function can detect that any operand will not return a single valid GeoJson object. Otherwise, the runtime behavior is as follows:  Returns false if any operand returns 0 or more than 1 items.  Returns NULL if any operand returns NULL.  Returns false if any operand returns an item that is not a valid GeoJson object.  Finally, if both operands return a single GeoJson object, returns true if the two geometries have any points in common. Otherwise, returns false.



boolea n	<ul> <li>Raises an error at compile time if the function can detect that any operand will not return a single valid GeoJson object. Otherwise, the runtime behavior is as follows:</li> <li>Returns false if any operand returns 0 or more than 1 item.</li> <li>Returns NULL if any operand returns NULL.</li> <li>Returns false if any operand returns an item that is not a valid GeoJson object.</li> <li>Finally, if both operands return a single GeoJson object and the second GeoJson is a polygon, the function returns true if the first geometry is completely contained inside the second polygon, with all of its points belonging to the interior of the polygon. The interior of a polygon is all the points in the polygon, except the points of the linear rings that define the polygon's boundary. Otherwise, returns false.</li> <li>Raises an error at compile time if the function detects that the first two operands will not return a single valid GeoJson object.</li> <li>Otherwise, the runtime behavior is as follows:</li> <li>Returns false if any of the first two operands returns 0 or more than 1 item.</li> <li>Returns NULL if any of the first two operands returns an item that is not a valid GeoJson object.</li> <li>Finally, if both of the first two operands return a single GeoJson object, the function returns true if the first geometry is within a distance of N meters from the second geometry, where N is the number returned by the third operand. The</li> </ul>
	<ul> <li>two operands will not return a single valid GeoJson object.</li> <li>Otherwise, the runtime behavior is as follows:</li> <li>Returns false if any of the first two operands returns 0 or more than 1 item.</li> <li>Returns NULL if any of the first two operands returns NULL.</li> <li>Returns false if any of the first two operands returns an item that is not a valid GeoJson object.</li> <li>Finally, if both of the first two operands return a single GeoJson object, the function returns true if the first geometry is within a distance of N meters from the second geometry,</li> </ul>
	distance between 2 geometries is defined as the minimum among the distances of any pair of points where the first point belongs to the first geometry, and the second point to the second geometry. Otherwise, returns false.
boolea n	The geo_near funcion is converted internally to a geo_within_distance function, with an an (implicit) order by the distance between the two geometries. However, if the query has an (explicit) order-by already, the function performs no ordering by distance. The geo_near function can appear only in the WHERE clause, and must be a top-level predicate. The geo_near function cannot be nested under an OR or NOT operator.
double boolea n	Raises an error at compile time if the function detects that an operand will not return a single valid GeoJson object. Otherwise, the runtime behavior is as follows:  Returns -1 if any of the operands returns zero or more than 1 item.  Returns NULL if any operand returns NULL.  Otherwise the function returns the geodetic distance between the 2 input geometries. The returned distance is the minimum among the distances of any pair of points, where the first point belongs to the first geometry and the second point to the second geometry. Between two such points, their distance is the length of the geodetic line that connects the points.  Returns false if an operand returns Zero or more than 1 item.  Returns NULL if an operand returns NULL.
1	double



6

# Working With Indexes

The SQL for Oracle NoSQL Database query processor can detect which of the existing indexes on a table can be used to optimize the execution of a query. This chapter provides a brief examples-based introduction to index creation, and queries using indexes. For a more detailed description of index creation and usage, see *SQL Reference Guide*.

To make it possible to fit the example output on the page, the examples in this chapter use mode LINE.

## **Basic Indexing**

This section builds on the examples that you began in Working with complex data.

```
sql-> mode LINE
Query output mode is LINE
sql-> create index idx income on Persons (income);
Statement completed successfully
sql-> create index idx age on Persons (age);
Statement completed successfully
sql-> SELECT * from Persons
WHERE income > 10000000 and age < 40;
> Row 0
| firstname | John
| lastname | Morgan
| age | 38
+----+
| income | 100000000
address
         street
                              | 187 Aspen Drive |
          | city
| state
                             | Middleburg |
                             | FL
                       | NULL
          | zipcode
          | phones
          | type
| areacode
| number
                             | work
                             1 305
                              | 1234079
                             | home
              type
              areacode
                             | 305
              number
                              2066401
```



connections	+	       
expenses	food   gas   travel	2000     10     700

1 row returned

## **Using Index Hints**

In the previous section, both indexes are applicable. For index idx\_income, the query condition income > 10000000 can be used as the starting point for an index scan that will retrieve only the index entries and associated table rows that satisfy this condition. Similarly, for index idx\_age, the condition age < 40 can be used as the stopping point for the index scan. SQL for Oracle NoSQL Database has no way of knowing which of the 2 predicates is more selective, and it assigns the same "value" to each index, eventually picking the one whose name is first alphabetically. In the previous example, idx\_age was used. To choose the idx\_income index instead, the query should be written with an index hint:

```
sql-> SELECT /*+ FORCE INDEX(Persons idx income) */ * from Persons
WHERE income > 10000000 and age < 40;
> Row 0
+----+
+-----+
| firstname | John
+-----
| lastname | Morgan
+-----
| income | 100000000
street
                      | 187 Aspen Drive |
       | city
                      | Middleburg |
        state
                     | FL
                  | NULL
        | zipcode
        phones
                     | work
          areacode
number
                     | 305
                      | 1234079
                     | home
          areacode
                     | 305
                      | 2066401
```



connections   	1   4   2		
+	food   gas   travel	2000   10   700	-+        -+

1 row returned

As shown above, hints are written as a special kind of comment that must be placed immediately after the SELECT keyword. What distinguishes a hint from a regular comment is the "+" character immediately after (without any space) the opening "/\*".

## **Complex Indexes**

The following example demonstrates indexing of multiple table fields, indexing of nested fields, and the use of "filtering" predicates during index scans.

```
sql-> create index idx_state_city_income on
Persons (address.state, address.city, income);
Statement completed successfully
sql-> SELECT * from Persons p WHERE p.address.state = "MA"
and income > 79000;
> Row 0
+-----
+----+
| income | 80000
| address
         | street
                             | 364 Mulberry Street |
          | city
          - = y
| state
                            | Leominster
                            | MA
          | zipcode
                            | NULL
          | phones
                             | work
              type
              type
areacode
                            | 339
              number
                            | 4120211
                            | work
              type
              areacode
                            | 339
              number
                             8694021
                             | home
              type
```



 	areacode   number 	339     1205678
     	type   areacode   number	home
connections	3   5   1   2	     
expenses	books   clothes   food   shoes	240

<sup>1</sup> row returned

Index idx\_state\_city\_income is applicable to the above query. Specifically, the state = "MA" condition can be used to establish the boundaries of the index scan (only index entries whose first field is "MA" will be scanned). Further, during the index scan, the income condition can be used as a "filtering" condition, to skip index entries whose third field is less or equal to 79000. As a result, only rows that satisfy both conditions are retrieved from the table.

## Multi-Key Indexes

A multi-key index indexes all the elements of an array, or all the elements and/or all the keys of a map. For such indexes, for each table row, the index contains as many entries as the number of elements/entries in the array/map that is being indexed. Only one array/map may be indexed.

```
sql-> create index idx_areacode on
Persons (address.phones[].areacode);
Statement completed successfully
sql-> SELECT * FROM Persons p WHERE
p.address.phones.areacode = any 339;
> Row 0
```

	+
id	2
firstname	
lastname	
age	
income	
lastLogin	2016-11-28T13:01:11.2088
T	T



address             	street   city   state   zipcode   phones   type   areacode   number	187 Hill Street
connections	1   3	
expenses	books   food   travel 	100
> Row 1		
id	4 +	
firstname	ı	i
lastname	ı	i
age	38	 
income	80000	 
ı	2016-10-19T09:18:05.5555	
address	street   city   state   zipcode   phones   type   areacode   number     type   areacode   number     type   areacode   number	364 Mulberry Street     Leominster     MA
connections	+   3   5   1	



	2		
expenses	books clothes food shoes	1	240 2000 6000 1200
> Row 2			
id			İ
firstname			İ
lastname			i
age			i
income	400000		 
lastLogin	2016-11-08T09:16:46.3929		 
address	street city state zipcode phones type areacode number  type areacode number  type areacode number		427 Linden Avenue   Monroe Township   NJ   NULL   work   201   3213267   work   201   8765421   home   339   3414578
+	2 4 1 3		
expenses	clothes food shoes		1500   900   1000

#### 3 rows returned

In the above example, a multi-key index is created on all the area codes in the Persons table, mapping each area code to the persons that have a phone number with that area code. The query is looking for persons who have a phone number with area

code 339. The index is applicable to the query and so the key 339 will be searched for in the index and all the associated table rows will be retrieved.

```
sql-> create index idx_expenses on
Persons (expenses.keys(), expenses.values());
Statement completed successfully
sql-> SELECT * FROM Persons p WHERE p.expenses.food > 1000;
> Row 0
| firstname | John
| lastname | Anderson
+----
    | 35
| income | 100000
address
        street
                         | 187 Hill Street |
                         | Beloit
        | city
         | state
                         | WI
         | zipcode
                   | 53511
         | phones
         type areacode
                         | home
                         | 339
                          | 1684972
            number
| connections | 1
    | 3
                | 100
| expenses | books
       | food
                         | 1700
                         | 2100
        | travel
| firstname | John
+-----
| lastname | Morgan
| income | 100000000
| address | street
                         | 187 Aspen Drive |
| city
                         | Middleburg |
```



	state   zipcode   phones   type   areacode   number   type   areacode   number	FL
	4   2 	   
· -	food   gas   travel 	2000     10     700
> Row 2	·	
id	4 	
firstname		
lastname		
age	38 <del> </del>	
income	80000 <del> </del>	
lastLogin	2016-10-19T09:18:05.5555	
address	street city state zipcode phones type areacode number  type areacode number  type areacode number  type areacode number	364 Mulberry Street     Leominster     MA
	number	8064321



connections	3   5   1   2		     
expenses	books   clothes   food   shoes	240   2000   6000   1200	

3 rows returned

In the above example, a multi-key index is created on all the expenses entries in the Persons table, mapping each category C and each amount A associated with that category to the persons that have an entry (C, A) in their expenses map. The query is looking for persons who spent more than 1000 on food. The index is applicable to the query and so only the index entries whose first field (the map key) is equal to "food" and second key (the amount) is greater than 1000 will be scanned and the associated rows retrieved.

## **Indexing JSON Data**

An index is a JSON index if it indexes at least one field that is contained inside JSON data.

Because JSON is schema-less, it is possible for JSON data to differ in type across table rows. However, when indexing JSON data, the data type must be consistent across table rows or the index creation will fail. Further, once one or more JSON indexes have been created, any attempt to write data of an incorrect type will fail.

With the exception of the previous restriction, indexing JSON data and working with JSON indexes behaves in much the same way as indexing non-JSON data. To create the index, specify a path to the JSON field using dot notation. You must also specify the data's type, using the AS keyword.

The following examples are built on the examples shown in Working with JSON.

```
sql-> create index idx_json_income on JSONPersons (person.income
as integer);
Statement completed successfully
sql-> create index idx_json_age on JSONPersons (person.age as integer);
Statement completed successfully
sql->
```

You can then run a query in the normal way, and the index idx\_json\_income will be automatically used. But as shown at the beginning of this chapter (Basic Indexing), the query processor will not know which index to use. To require the use of a particular index provide an index hint as normal:



id	3	
person	address city phones areacode number type	Middleburg
	areacode number type state street age connections	305
	expenses food gas travel firstname income lastLogin lastname	2000

1 row returned
sql->

Finally, when creating a multi-key index on a JSON map, a type must not be given for the .keys() expression. This is because the type will always be String. However, a type declaration is required for the .values() expression:



	lastLogin	35
> Row 1 +	 3	
+	address	
	phones     areacode     number     type     areacode     number     type     state     street age connections  expenses     food     gas     travel firstname income	1234079
+		+
id	4	
person	address city phones areacode number type	Leominster



areacode number type  areacode number type	8694021   work
areacode number type state street age connections	305
expenses books clothes food shoes firstname income lastLogin lastname	240

3 rows returned
sql->

Be aware that all the other constraints that apply to a non-JSON multi-keyed index also apply to a JSON multi-keyed index.



7

# Working with Table Rows

This chapter provides examples on how to insert and update table rows using SQL for Oracle NoSQL Database INSERT and UPDATE statements.

## Adding Table Rows using INSERT and UPSERT

This topic provides examples on how to add table rows using the SQL for Oracle NoSQL Database INSERT and UPSERT statements.

You use the INSERT statement to insert or update a single row in an existing table.

### **Examples:**

If you executed the SQLBasicExamples Script, you should already have created the table named Users. The table had this definition:

```
CREATE TABLE Users
 id integer,
 firstname string,
 lastname string,
 age integer,
 income integer,
 primary key (id)
sql-> describe table Users;
=== Information ===
+----+
| name | ttl | owner | sysTable | r2compat | parent | children | indexes |
description |
+----+
          | N | N |
| Users | |
+----+
=== Fields ===
+---+
| id | name | type | nullable | default | shardKey | primaryKey |
identity |
+---+-----
+----+
| 1 | id
      | Integer | N | NullValue | Y | Y
```

```
+----+
+---+-----
+----+
1
+---+-----
+----+
| 4 | age | Integer | Y | NullValue |
  +----+
| 5 | income | Integer | Y | NullValue |
+---+-----
+----+
```

To insert a new row into the Users table, use the INSERT statement as follows. Because you are adding values to all table columns, you do not need to specify column names explicitly:

```
sql-> INSERT INTO Users VALUES (10, "John", "Smith", 22, 45000);
{"NumRowsInserted":1}
1 row returned
sql-> select * from Users;
{"id":10,"firstname":"John","lastname":"Smith","age":22,"income":45000}
```

To insert data into some, but not all, table columns, specify the column names explicitly in the INSERT statement. Any columns that you do not specify are assigned either <code>NULL</code> or the default value supplied when you created the table:

```
sql-> INSERT INTO Users (id, firstname, income)
VALUES (11, "Mary", 5000);
{"NumRowsInserted":1}
1 row returned

sql-> select * from Users;
{"id":11,"firstname":"Mary","lastname":null,"age":null,"income":5000}
{"id":10,"firstname":"John","lastname":"Smith","age":22,"income":45000}
2 rows returned
```

#### **Using the UPSERT Statement**

The word UPSERT combines UPDATE and INSERT, describing it statement's function. Use an UPSERT statement to insert a row where it does not exist, or to update the row with new values when it does.

For example, if you already inserted a new row as described in the previous section, executing the next statement *updates* user John's age to 27, and income to 60,000. If



you did not execute the previous INSERT statement, the UPSERT statement inserts a new row with user id 10 to the Users table.

```
sql-> UPSERT INTO Users VALUES (10, "John", "Smith", 27, 60000);
{"NumRowsInserted":0}
1 row returned
sql-> UPSERT INTO Users VALUES (11, "Mary", "Brown", 28, 70000);
{"NumRowsInserted":0}
1 row returned
sql-> select * from Users;
{"id":10,"firstname":"John","lastname":"Smith","age":22,"income":60000}
{"id":11,"firstname":"Mary","lastname":"Brown","age":28,"income":70000}
2 rows returned
```

#### **Using an IDENTITY Column**

You can use IDENTITY columns to automatically generate values for a table column each time you insert a new table row. See Identity Column in the *SQL Reference Guide*.

Here are a few examples for how to use the INSERT statements for both flavors of an IDENTITY column:

- GENERATED ALWAYS AS IDENTITY
- GENERATED BY DEFAULT [ON NULL] AS IDENTITY

Create a table named <code>Employee\_test</code> using one column, DeptId, as GENERATED ALWAYS AS IDENTITY. This IDENTITY column is not the primary key. Insert a few rows into the table.

```
sql-> CREATE TABLE EmployeeTest
(
    Empl_id INTEGER,
    Name STRING,
    DeptId INTEGER GENERATED ALWAYS AS IDENTITY (CACHE 1),
    PRIMARY KEY(Empl_id)
);

INSERT INTO Employee_test VALUES (148, 'Sally', DEFAULT);
INSERT INTO Employee_test VALUES (250, 'Joe', DEFAULT);
INSERT INTO Employee test VALUES (346, 'Dave', DEFAULT);
```

The INSERT statement inserts the following rows with the system generates values 1, 2, and 3 for the IDENTITY column <code>DeptId</code>.

Empl_id	Name	Deptld
148	Sally	1
250	Joe	2
346	Dave	3

You cannot specify a value for the <code>DeptId</code> IDENTITY column when inserting a row to the <code>Employee\_test</code> table, because you defined that column as <code>GENERATED ALWAYS AS IDENTITY</code>. Specifying <code>DEFAULT</code> as the column value, the system generates the next <code>IDENTITY</code> value.



Conversely, trying to execute the following SQL statement causes an exception, because you supply a value (200) for the Dept1d column.

```
sql-> INSERT INTO Employee test VALUES (566, 'Jane', 200);
```

If you create the column as <code>GENERATED</code> BY <code>DEFAULT</code> AS <code>IDENTITY</code> for the <code>Employee\_test</code> table, the system generates a value only if you fail to supply one. For example, if you define the <code>Employee\_test</code> table as follows, then execute the <code>INSERT</code> statement as above, the statement inserts the value 200 for the employee's <code>DeptId</code> column.

```
CREATE Table Employee_test
(
    Empl_id INTEGER,
    Name STRING,
    Deptid INTEGER GENERATED BY DEFAULT AS IDENTITY (CACHE 1),
    PRIMARY KEY(Empl_id)
);
```

# Modifying Table Rows using UPDATE Statements

This topic provides examples of how to update table rows using SQL for Oracle NoSQL Database UPDATE statements. These are an efficient way to update table row data, because UPDATE statements make *server-side updates* directly, without requiring a Read/Modify/Write update cycle.



You can use UPDATE statements to update only an existing row. You cannot use UPDATE to either create new rows, or delete existing rows. An UPDATE statement can modify only a single row at a time.

# **Example Data**

This chapter's examples uses the data loaded by the SQLJSONExamples script, which can be found in the Examples download package. For details on using this script, the sample data it loads, and the Examples download, see See SQLJSONExamples Script.

# Changing Field Values

In the simplest case, you can change the value of a field using the Update Statement SET clause. The JSON example data set has a row which contains just an array and an integer. This is row ID 6:

```
sql-> mode column
Query output mode is COLUMN
sql-> SELECT * from JSONPersons j WHERE j.id = 6;
+---+
```

1	id	person
+	 6	   myarray
		1 1
		2
		3
		4
		mynumber   5
+		+

1 row returned

1 row returned

You can change the value of mynumber in that row using the following statement:

In the previous example, the results returned by the Update statement was not very informative, so we were required to reissue the Select statement in order to view the results of the update. You can avoid that by using a RETURNING clause. This functions exactly like a Select statement:



You can further limit and customize the displayed results in the same way that you can do so using a SELECT statement:

It is normally possible to update the value of a non-JSON field using the SET clause. However, you cannot change a field if it is a primary key. For example:

## Modifying Array Values

You use the Update statement ADD clause to add elements into an array. You use a SET clause to change the value of an existing array element. And you use a REMOVE clause to remove elements from an array.

# Adding Elements to an Array

The ADD clause requires you to identify the array position that you want to operate on, followed by the value you want to set to that position in the array. If the index value that you set is 0 or a negative number, the value that you specify is inserted at the beginning of the array.



If you do not provide an index position, the array value that you specify is appended to the end of the array.

```
sql-> SELECT * from JSONPersons j WHERE j.id = 6;
+---+
| id | person
+---+
| mynumber | 300 |
+---+
1 row returned
sql-> UPDATE JSONPersons j
     ADD j.person.myarray 0 50,
     ADD j.person.myarray 100
     WHERE j.id = 6
     RETURNING *;
+---+
| id | person |
3
          4
  100
  | mynumber | 300 |
1 row returned
sql->
```

Notice that multiple ADD clauses are used in the guery above.

Array values get appended to the end of the array, even if you provide an array position that is larger than the size of the array. You can either provide an arbitrarily large number, or make use of the size() function:



You can append values to the array using the built-in seq concat () function:

```
sql-> UPDATE JSONPersons j
       ADD j.person.myarray seq_concat(66, 77, 88)
       WHERE j.id = 6
       RETURNING *;
+---+
| id | person |
+---+
  6 | myarray
| 6 | myarray | 50 | 1 | 1 |
             2
            3
             4
             100 |
            400
            66 I
             77
    | mynumber | 300
1 row returned
sql->
```

If you provide an array position that is between 0 and the array's size, then the value you specify will be inserted into the array *before* the specified position. To determine the correct position, start counting from 0:



## Changing an Existing Element in an Array

To change an existing value in an array, use the SET clause and identify the value's position using []. To determine the value's position, start counting from 0:

```
sql-> UPDATE JSONPersons j
      SET j.person.myarray[3] = 1000
      WHERE j.id = 6
      RETURNING *;
+---+
       person |
| id |
+---+
1000 |
           3
           100 |
           400
           77
           88
   | mynumber | 300 |
1 row returned
sql->
```

## Removing Elements from Arrays

To remove an existing element from an array, use the REMOVE clause. To do this, you must identify the position of the element in the array that you want to remove. To determine the value's position, start counting from 0:



+-	+		+	F
	6	myarray		
	-		50	
			1	
			2	
			3	
			4	
			100	
			400	
			66	
			77	
			88	
		mynumber	300	
+-	+		+	H

It is possible for the array position to be identified by an expression. For example, in our sample data, some records include an array of phone numbers, and some of those phone numbers include a work number:

```
sql-> SELECT * FROM JSONPersons j WHERE j.id = 3;
+---+
                person
| 3 | address
 | city | Middleburg
| phones
      areacode | 305
         number | 1234079
         type | work
         areacode | 305
         number | 2066401
      type | home
state | FL
street | 187 Aspen Drive
               | 38
   | age
   | connections
                 1
                 4
                 2
    expenses
   | food | 2000
| gas | 10
   | travel
               1 700
```



```
1 row returned
sql->
```

We can remove the work number from the array in one of two ways. First, we can directly specify its position in the array (position 0), but that only removes a single element at a time. If we want to remove all the work numbers, we can do it by using the \$element variable. To illustrate, we first add another work number to the array:

```
sql-> UPDATE JSONPersons j
        ADD j.person.address.phones 0
        {"type":"work", "areacode":415, "number":9998877}
        WHERE j.id = 3
        RETURNING *;
          person
  3 | address
    | city | Middleburg
    phones
         areacode | 415
            number | 9998877
    | number | 99988
| type | work
           areacode | 305
            number | 1234079
             type | work
            areacode | 305
            number | 2066401
             type | home
          state | FL
street | 187 Aspen Drive
                    | 38
     age
     | connections
                      1
                      2
     | expenses
                   | 2000
     | food
         gas
                    | 10
    | travel | 700
| firstname | John
| income | 100000000
                    | 2016-11-29T08:21:35.4971 |
     | lastLogin
     | lastname
                    | Morgan
1 row returned
sql->
```

Now we can remove all the work numbers as follows:



```
WHERE j.id = 3
      RETURNING *;
| 3 | address
  | city | Middleburg
      phones
       areacode | 305
  | number | 2066401
| type | home
| state | FL
| street | 187 Aspen Drive
| age | 38
   | connections
                   1
                    4
  expenses
```

## Modifying Map Values

To write a new field to a map, use the PUT clause. You can also use the PUT clause to change an existing map value. To remove a map field, use the REMOVE clause.

For example, consider the following two rows from our sample data:

```
sql-> SELECT * FROM JSONPersons j WHERE j.id = 6 OR j.id = 3;
+---+
                   person
| 3 | address
   | city | Middleburg
    | phones
    | areacode | 305
| number | 2066401
| type | home
| state | FL
| street | 187 Aspen Drive
| age | 38
     | connections
                       1
                       4
```



```
expenses
               | 2000
   l food
  | gas | 10
| travel | 700
| firstname | John
                | 10000000
   | income
   | lastname
                | Morgan
  6 | myarray
                  50
                   1
                   2
                   3
                   4
                  100
                   400
                   66
                   77
                  88
   | mynumber | 300
2 rows returned
sql->
```

These two rows look nothing alike. Row 3 contains information about a person, while row 6 contains, essentially, random data. This is possible because the person column is of type JSON, which is not strongly typed. But because we interact with JSON columns as if they are maps, we can fix row 6 by modifying it as a map.

## Removing Elements from a Map

To begin, we remove the two existing elements from row six (myarray and mynumber). We do this with a single UPDATE statement, which allows us to execute multiple update clauses so long as they are comma-separated:



## Adding Elements to a Map

Next, we add person data to this table row. We could do this with a single UPDATE statement by specifying the entire map with a single PUT clause, but for illustration purposes we do this in multiple steps.

To begin, we specify the person's name. Here, we use a single PUT clause that specifies a map with multiple elements:

Next, we specify the age, connections, expenses, income, and lastLogin fields using multiple PUT clauses on a single UPDATE statement:

```
sql-> UPDATE JSONPersons j
        PUT j.person {"age" : 43},
        PUT j.person {"connections" : [2,3]},
        PUT j.person {"expenses" : {"food" : 1100,
                              "books" : 210,
                              "travel" : 50}},
        PUT j.person {"income" : 80000},
        PUT j.person {"lastLogin" : "2017-06-29T16:12:35.0285"}
        WHERE j.id = 6
       RETURNING *;
                 person
 | id |
  6 | age | 43
   | connections
                3
   expenses
    | books | 210
        food | 1100
    | travel | 50
   | firstname | Wendy
    | income | 80000
    | lastname | Purvis
```



```
1 row returned
sql->
```

We still need an address. Again, we could do this with a single PUT clause, but for illustration purposes we will use multiple clauses. Our first PUT creates the address element, which uses a map as a value. Our second PUT adds elements to the address map:

```
sql-> UPDATE JSONPersons j
     PUT j.person {"address" : {"street" : "479 South Way Dr"}},
     PUT j.person.address {"city" : "St. Petersburg",
                     "state" : "FL"}
     WHERE j.id = 6
     RETURNING *;
         person
| 6 | address
   | city | St. Petersburg
       state | FL
   | street | 479 South Way Dr
   | age | 43
   | connections
               3
   expenses
   | books | 210
       food | 1100
    | travel | 50
   | firstname | Wendy
   | income | 80000
   | lastname | Purvis
1 row returned
sql->
```

Finally, we provide phone numbers for this person. These are specified as an array of maps:



```
phones
     areacode | 727
       number | 8284321
       type | work
      areacode | 727
       number | 5710076
        type | home
        areacode | 727
        number | 8913080
     type | mobile
state | FL
street | 479 South Way Dr
| age
               | 43
| connections
                 3
expenses
| books | 210
| food | 1100
| travel | 50
| firstname | Wendy
```

## **Updating Existing Map Elements**

To update an existing element in a map, you can use the PUT clause in exactly the same way as you add a new element to map. For example, to update the lastLogin time:

#### Alternatively, use a SET clause:

```
sql-> UPDATE JSONPersons j
       SET j.person.lastLogin = "2017-06-29T20:38:56.2751"
       WHERE j.id = 6
       RETURNING *;
                  person
+---+
 | 6 | address
   | city | St. Petersburg
   phones
        areacode | 727
           number | 8284321
          type | work
          areacode | 727
           number | 5710076
           type | home
         areacode | 727
         number | 8913080
type | mobile
       state | FL
street | 479 South Way Dr
    | age
    | connections
                   3
    expenses
               | 210
    books
```



```
| | food | 1100 | | travel | 50 | | | | firstname | Wendy | | | income | 80000 | | | lastLogin | 2017-06-29T20:38:56.2751 | | lastname | Purvis | |
```

If you want to set the timestamp to the current time, use the <code>current\_time()</code> built-in function.

```
sql-> UPDATE JSONPersons j
       SET j.person.lastLogin = cast(current time() AS String)
       WHERE j.id = 6
       RETURNING *;
+---+
| 6 | address
 city | St. Petersburg
   | phones
       areacode | 727
          number | 8284321
          type | work
        areacode | 727
number | 5710076
          type | home
         areacode | 727
          number | 8913080
    type | mobile | state | FL | street | 479 South Way Dr
                  | 43
    age
    | connections
                   2
                   3
   | expenses
   | books | 210
| food | 1100
    | travel | 50
| firstname | Wendy
```

1 row returned
sql->



If an element in the map is an array, you can modify it in the same way as you would any array. For example:

```
sql-> UPDATE JSONPersons j
       ADD j.person.connections seq concat(1, 4)
       WHERE j.id = 6
       RETURNING *;
+---+
| id |
                  person
| 6 | address
   city | St. Petersburg
   | phones
       areacode | 727
          number | 8284321
           type | work
        areacode | 727
number | 5710076
          type | home
          areacode | 727
          number | 8913080
    type | mobile | state | FL | street | 479 South Way Dr | age | 43
    connections
                    3
                    1
   expenses
    | books | 210
| food | 1100
       travel
                 | 50
```

1 row returned

If you are unsure of an element being an array or a map, you can use both ADD and PUT within the same UPDATE statement. For example:



++			+
6	address		1
1 1	city		St. Petersburg
	phones		1
1 1	areacode		727
	number		8284321
	type		work
	areacode	ı	727
1 1	number	1	5710076
1 1	type	i	
i i	cype	1	
i i	areacode	1	727
1 1	number		8913080
1 1	type		mobile
1 1	state		FL
1 1	street		479 South Way Dr
	age		43
	connections		
			2   3
			1
			4
1 1			5
			7
i i	expenses		i
1 1	books		210
	food		1100
1 1	travel		50
1 1	firstname		Wendy
1 1	income		80000
	lastLogin		2017-06-29T04:40:15.917
	lastname		Purvis
++			+

1 row returned

If the element is an array, the ADD gets applied and the PUT is a noop. If it is a map, then the PUT gets applied and ADD is a noop. In this example, since the element is an array, the ADD gets applied.

# Managing Time to Live Values

Time to Live (TTL) values indicate how long data can exist in a table before it expires. Expired data can no longer be returned as part of a query.

Default TTL values can be set on either a table-level or a row level when the table is first defined. Using UPDATE statements, you can change the TTL value for a single row.

You can see a row's TTL value using the  $remaining_hours()$ ,  $remaining_days()$  or expiration time() built-in functions. These TTL functions require a row as input. We

accomplish this by using the \$ as part of the table alias. This causes the table alias to function as a row variable.

```
sql-> SELECT remaining_days($j) AS Expires
     FROM JSONPersons $j WHERE id = 6;
+-----+
| Expires |
+-----+
| -1 |
+----+
1 row returned
sql->
```

The previous query returns -1. This means that the row has no expiration time. We can specify an expiration time for the row by using an UPDATE statement with a set <code>TTL</code> clause. This clause computes a new TTL by specifying an offset from the current expiration time. If the row never expires, then the current expiration time is <code>1970-01-01T00:00:00:00.000</code>. The value you provide to <code>set TTL</code> must specify units of either <code>HOURS</code> or <code>DAYS</code>.

To see the new expiration time, we can use the built-in <code>expiration\_time()</code> function. Because we specified an expiration time based on a day boundary, the row expires at midnight of the following day (expiration rounds up):



To turn off the TTL so that the row will never expire, specify a negative value, using either HOURS or DAYS as the unit:

Notice that the RETURNING clause provides a value of 0 days. This indicates that the row will never expire. Further, if we look at the remaining\_days() using a SELECT statement, we will once again see a negative value, indicating that the row never expires:

```
sql-> SELECT remaining_days($j) AS Expires
    FROM JSONPersons $j WHERE id = 6;
+-----+
| Expires |
+-----+
| -1 |
+-----+
1 row returned
sql->
```

## Avoiding the Read-Modify-Write Cycle

An important aspect of UPDATE Statements is that you do not have to read a value in order to update it. Instead, you can blindly modify a value directly in the store without ever retrieving (reading) it. To do this, you refer to the value you want to modify using the \$ variable.

For example, we have a row in JSONPersons that looks like this:



1 row returned

We can blindly update the value of the person.expenses.books field by referencing \$. In the following statement, no read is performed on the store. Instead, the write operation is performed directly at the store.

1 row returned

To see that the books expenses value has indeed been incremented by 100, we perform a second  $\mathtt{SELECT}$  statement.



		727   5710076   home
		8913080   mobile   FL   479 South Way Dr
age connections	 	43
expenses books food travel firstname income lastLogin lastname	         	310   1100   50   Wendy   80000   2017-07-25T22:50:06.482   Purvis

1 row returned



A

# Introduction to the SQL for Oracle NoSQL Database Shell

This appendix describes how to configure, start and use the SQL for Oracle NoSQL Database shell to execute SQL statements. This section also describes the available shell commands.

You can directly execute DDL, DML, user management, security, and informational statements using the SQL shell.

# Running the SQL Shell

You can run the SQL shell interactively or use it to run single commands. Here is the general usage to start the shell:

#### where:

-consistency

Configures the read consistency used for this session.

-durability

Configures the write durability used for this session.

-helper-hosts

Specifies a comma-separated list of hosts and ports.

-store

Specifies the name of the store.

-timeout

Configures the request timeout used for this session.

username

Specifies the username to login as.

For example, you can start the shell like this:

```
java -jar KVHOME/lib/sql.jar
-helper-hosts node01:5000 -store kvstore
sql->
```

This command assumes that a store kystore is running at port 5000. After the SQL starts successfully, you execute queries. In the next part of this document, you will find an introduction to SQL for Oracle NoSQL Database and how to create query statements.

If you want to import records from a file in either JSON or CSV format, you can use the import command. For more information see import.

If you want to run a script, use the load command. For more information see load.

For a complete list of utility commands accessed through "java -jar" <kvhome>/lib/sql.jar <command> see Shell Utility Commands.

# Configuring the shell

You can also set the shell start-up arguments by modifying the configuration file . ${\tt kvclirc}$  found in your home directory.

Arguments can be configured in the .kvclirc file using the name=value format. This file is shared by all shells, each having its named section. [sql] is used for the Query shell, while [kvcli] is used for the Admin Command Line Interface (CLI).

For example, the .kvclirc file would then contain content like this:

```
helper-hosts=node01:5000
store=kvstore
timeout=10000
consistency=NONE REQUIRED
durability=COMMIT_NO_SYNC
username=root
security=/tmp/login root
[kvcli]
host=node01
port=5000
store=kvstore
admin-host=node01
admin-port=5001
username=user1
security=/tmp/login user
admin-username=root
admin-security=/tmp/login root
timeout=10000
consistency=NONE REQUIRED
durability=COMMIT NO SYNC
```



# **Shell Utility Commands**

The following sections describe the utility commands accessed through "java -jar" <kvhome>/lib/sql.jar <command>".

The interactive prompt for the shell is:

```
sql->
```

The shell comprises a number of commands. All commands accept the following flags:

-help

Displays online help for the command.

• ?

Synonymous with -help. Displays online help for the command.

The shell commands have the following general format:

1. All commands are structured like this:

```
sql-> command [arguments]
```

- 2. All arguments are specified using flags that start with "-"
- 3. Commands and subcommands are case-insensitive and match on partial strings(prefixes) if possible. The arguments, however, are case-sensitive.

#### connect

Connects to a KVStore to perform data access functions. If the instance is secured, you may need to provide login credentials.

## consistency

```
consistency [[NONE_REQUIRED | NONE_REQUIRED_NO_MASTER |
ABSOLUTE] [-time -permissible-lag <time_ms> -timeout <time_ms>]]
```

Configures the read consistency used for this session.



## describe

```
describe | desc [as json]
  {table table_name [field_name[,...] ] |
  index index_name on table_name
  }
```

Describes information about a table or index, optionally in JSON format.

Specify a fully-qualified table name as follows:

Entry specification	Description
table_name	Required. Specifies the full table name. Without further qualification, this entry indicates a table created in the default namespace (sysdefault), which you do not have to specify.
parent-table.child-table	Specifies a child table of a parent. Specify the parent table followed by a period (.) before the child name. For example, if the parent table is Users, specify the child table named MailingAddress as Users. MailingAddress.
namespace-name:table-name	Specifies a table created in the non-default namespace. Use the namespace followed by a colon (:). For example, to reference table <code>Users</code> , created in the <code>Sales</code> namespace, enter <code>table_name</code> as <code>Sales:Users</code> .

Following is the output of describe for table ns1:t1:

```
sql-> describe table ns1:t1;
=== Information ===
+----
+----+
| namespace | name | ttl | owner | sysTable | r2compat | parent |
children | indexes | description |
+----
+----+
| N |
+-----
+----+
=== Fields ===
+----+
| id | name | type | nullable | default | shardKey | primaryKey
| identity |
```



This example shows using describe as json for the same table:

```
sql-> describe as json table ns1:t1;
  "json_version" : 1,
  "type" : "table",
 "name" : "t1",
  "namespace" : "ns1",
  "shardKey" : [ "id" ],
  "primaryKey" : [ "id" ],
  "fields" : [ {
    "name" : "id",
    "type" : "INTEGER",
    "nullable" : false,
    "default" : null
  }, {
    "name" : "name",
    "type" : "STRING",
    "nullable" : true,
    "default" : null
  } ]
}
```

# durability

```
durability [[COMMIT_WRITE_NO_SYNC | COMMIT_SYNC |
COMMIT_NO_SYNC] | [-master-sync <sync-policy> -replica-sync <sync-policy> -replica-ask <ack-policy>]] <sync-policy>: SYNC, NO_SYNC, WRITE_NO_SYNC <ack-policy>: ALL, NONE, SIMPLE MAJORITY
```

Configures the write durability used for this session.

# exit

exit | quit

Exits the interactive command shell.

# help

help [command]

Displays help message for all shell commands and sql command.

# history

```
history [-last < n>] [-from < n>] [-to < n>]
```

Displays command history. By default all history is displayed. Optional flags are used to choose ranges for display.

# import

```
import -table table name -file file name [JSON | CSV]
```

Imports records from the specified file into table table name.

Specify a fully-qualified table name as follows:

Entry specification	Description
table_name	Required. Specifies the full table name. Without further qualification, this entry indicates a table created in the default namespace (sysdefault), which you do not have to specify.
parent-table.child-table	Specifies a child table of a parent. Specify the parent table followed by a period (.) before the child name. For example, if the parent table is Users, specify the child table named MailingAddress as Users.MailingAddress.
namespace-name:table-name	Specifies a table created in the non-default namespace. Use the namespace followed by a colon (:). For example, to reference table <code>Users</code> , <code>created</code> in the <code>Sales</code> namespace, enter <code>table_name</code> as <code>Sales:Users</code> .

Use -table to specify the name of a table into which the records are loaded. The alternative way to specify the table is to add the table specification "Table: table\_name" before its records in the file.

For example, this file contains the records to insert into two tables, users and email:

Table: users
<records of users>
...
Table: emails
<record of emails>



The imported records can be either in JSON or CSV format. If you do not specify the format, JSON is assumed.

#### load

```
load -file <path to file>
```

Load the named file and interpret its contents as a script of commands to be executed. If any command in the script fails execution will end.

For example, suppose the following commands are collected in the script file test.sql:

```
### Begin Script ###
load -file test.ddl
import -table users -file users.json
### End Script ###
```

Where the file test.ddl would contain content like this:

```
DROP TABLE IF EXISTS users;

CREATE TABLE users(id INTEGER, firstname STRING, lastname STRING, age INTEGER, primary key (id));
```

And the file users.json would contain content like this:

```
{"id":1,"firstname":"Dean","lastname":"Morrison","age":51}
{"id":2,"firstname":"Idona","lastname":"Roman","age":36}
{"id":3,"firstname":"Bruno","lastname":"Nunez","age":49}
```

Then, the script can be run by using the load command in the shell:

```
> java -jar KVHOME/lib/sql.jar -helper-hosts node01:5000 \
-store kvstore
sql-> load -file ./test.sql
Statement completed successfully.
Statement completed successfully.
Loaded 3 rows to users.
```

## mode

```
mode [COLUMN | LINE | JSON [-pretty] | CSV]
```

Sets the output mode of query results. The default value is JSON.

For example, a table shown in COLUMN mode:

```
sql-> mode column;
sql-> SELECT * from users;
+----+
| id | firstname | lastname | age |
```



+-		-+		-+-		+-		+
-	8		Len		Aguirre		42	
	10		Montana		Maldonado		40	
	24		Chandler		Oneal		25	
	30		Pascale		Mcdonald		35	
	34		Xanthus		Jensen		55	
	35		Ursula		Dudley		32	
	39		Alan		Chang		40	
	6		Lionel		Church		30	
	25		Alyssa		Guerrero		43	
	33		Gannon		Bray		24	
	48		Ramona		Bass		43	
	76		Maxwell		Mcleod		26	
	82		Regina		Tillman		58	
	96		Iola		Herring		31	
	100		Keane		Sherman		23	
+-		-+		-+-		+-		-+

• • •

100 rows returned

#### Empty strings are displayed as an empty cell.

```
sql-> mode column;
sql-> SELECT * from tab1 where id = 1;
+---+---+
| id | s1 | s2 | s3 |
+---+---+
| 1 | NULL | | NULL |
+---+----+
```

1 row returned

#### For nested tables, identation is used to indicate the nesting under column mode:

```
sql-> SELECT * from nested;
+----+
+----+
| id | name |
details
                 | 1 | one |
address
| | city |
Waitakere
| country | French
Guiana
                    zipcode |
7229
attributes
color |
blue
```



	   	price size phone	   	expensive large [(08)2435-0742,	(09)8083-8862,	(08) 0742-2526]
+       +   +   +	3   three               	size		Viddalba Bhutan 280071 blue cheap small [(08)5361-2051,	(03)5502-9721,	(09) 7962-8693]

For example, a table shown in LINE mode, where the result is displayed vertically and one value is shown per line:

```
sql-> mode line;
sql-> SELECT * from users;
> Row 1
+----+
| firstname | Len
| lastname | Aguirre |
| age | 42 |
+----+
 > Row 2
+----+
| id | 10 |
| firstname | Montana |
| lastname | Maldonado |
+----+
 > Row 3
+----+
| id | 24 |
| firstname | Chandler |
| lastname | Oneal |
| age | 25
+----+
100 rows returned
```

Just as in COLUMN mode, empty strings are displayed as an empty cell:

```
sql-> mode line;
sql-> SELECT * from tabl where id = 1;
> Row 1
```



1 row returned

#### For example, a table shown in JSON mode:

```
sql-> mode ison;
sql-> SELECT * from users;
{"id":8,"firstname":"Len","lastname":"Aguirre","age":42}
{"id":10,"firstname":"Montana","lastname":"Maldonado","age":40}
{"id":24, "firstname": "Chandler", "lastname": "Oneal", "age":25}
{"id":30,"firstname":"Pascale","lastname":"Mcdonald","age":35}
{"id":34,"firstname":"Xanthus","lastname":"Jensen","age":55}
{"id":35,"firstname":"Ursula","lastname":"Dudley","age":32}
{"id":39,"firstname":"Alan","lastname":"Chang","age":40}
{"id":6,"firstname":"Lionel","lastname":"Church","age":30}
{"id":25, "firstname": "Alyssa", "lastname": "Guerrero", "age":43}
{"id":33,"firstname":"Gannon","lastname":"Bray","age":24}
{"id":48, "firstname": "Ramona", "lastname": "Bass", "age":43}
{"id":76, "firstname": "Maxwell", "lastname": "Mcleod", "age":26}
{"id":82,"firstname":"Regina","lastname":"Tillman","age":58}
{"id":96,"firstname":"Iola","lastname":"Herring","age":31}
{"id":100,"firstname":"Keane","lastname":"Sherman","age":23}
{"id":3,"firstname":"Bruno","lastname":"Nunez","age":49}
{"id":14,"firstname":"Thomas","lastname":"Wallace","age":48}
{"id":41,"firstname":"Vivien","lastname":"Hahn","age":47}
100 rows returned
```

#### Empty strings are displayed as "".

```
sql-> mode json;
sql-> SELECT * from tab1 where id = 1;
{"id":1,"s1":null,"s2":"","s3":"NULL"}
1 row returned
```

#### Finally, a table shown in CSV mode:

```
sql-> mode csv;
sql-> SELECT * from users;
8,Len,Aguirre,42
10,Montana,Maldonado,40
24,Chandler,Oneal,25
30,Pascale,Mcdonald,35
34,Xanthus,Jensen,55
35,Ursula,Dudley,32
39,Alan,Chang,40
```



```
6, Lionel, Church, 30
25, Alyssa, Guerrero, 43
33, Gannon, Bray, 24
48, Ramona, Bass, 43
76, Maxwell, Mcleod, 26
82, Regina, Tillman, 58
96, Iola, Herring, 31
100, Keane, Sherman, 23
3, Bruno, Nunez, 49
14, Thomas, Wallace, 48
41, Vivien, Hahn, 47
...
100 rows returned
```

Like in JSON mode, empty strings are displayed as "".

```
sql-> mode csv;
sql-> SELECT * from tab1 where id = 1;
1,NULL,"","NULL"

1 row returned
```



Only rows that contain simple type values can be displayed in CSV format. Nested values are not supported.

## output

```
output [stdout | file]
```

Enables or disables output of query results to a file. If no argument is specified, it shows the current output.

## page

```
page [on | <n> | off]
```

Turns query output paging on or off. If specified, n is used as the page height.

If n is 0, or "on" is specified, the default page height is used. Setting n to "off" turns paging off.

## show faults

```
show faults [-last] [-command <index>]
```

Encapsulates commands that display the state of the store and its components.

#### show indexes

```
show_indexes_statement ::= SHOW [AS JSON] INDEXES ON table_name
```

The show indexes statement provides the list of indexes present on a specified table. The parameter AS JSON is optional and can be specified if you want the output to be in JSON format.

#### Example 1: List indexes on the specified table

The following statement lists the indexes present on the users2 table.

```
SHOW INDEXES ON users2; indexes idx1
```

#### Example 2: List indexes on the specified table in JSON format

The following statement lists the indexes present on the users2 table in JSON format.

# show mrtable-agent-statistics

```
show mrtable-agent-statistics [-agent <agentID>][-table <tableName>][-
json]
```

Shows the latest statistics as of the last one minute for multi-region table agents. With no arguments, this command shows combined statistics over all regions the MR Table spans.

#### **Input Parameters**

Optionally, you can enable the following flags with appropriate parameters with this command:

**Table A-1** Input Parameters

Flag	Parameter	Description
- agent	agentID	Limits the statistics to the agent ID specified. You can find the agent ID from the JSON config file created while configuring your agent. See Configure XRegion Service.
- table	tableName	Limits the statistics to the MR Table specified.



Table A-1 (Cont.) Input Parameters

Flag	Parameter	Description
- json	<u>-</u>	Returns the complete statistics in JSON format. Even though the statistics are returned in JSON format by default, specifying this flag adds additional information in the output such as operation, return code, and the return code's description.

#### **Output Statistics**

The statistics reported by the show mrtable-agent-statistics can be categorized as those used to:

#### Monitor streams from other regions

Table A-2 Output Statistics 1

Statistic	Description
completeWriteOps	Number of complete write operations per region.
lastMessageMs	Timestamp when the agent sees the last message from a remote region, in milliseconds.
	If this statistic information is not available, -1 is printed as its output value.
lastModificationMs	Timestamp of the last operation performed in each remote region, in milliseconds.
	If this statistic information is not available, -1 is printed as its output value.
laggingMs (avg, max, min)	In a multi-region KVStore, each shard in a region pushes all the operations performed on all its tables to the agent's queue. The agent replicates the contents of its queue, in event order, to all other regions. The lagging statistic represents the time difference between an event being pushed into the queue and replicated to the other regions by the agent. If this value is high, it indicates that the queue is getting backed up. A smaller value indicates that the agent is able to keep up with the number of events coming from remote regions. The lagging statistics are reported as average, minimum, and maximum in milliseconds for each remote region.  If this statistic information is not available, -1 is
	G



Table A-2 (Cont.) Output Statistics 1

Statistic	Description
latencyMs (avg, max, min)	In MR tables, the latency statistic indicates the time taken in milliseconds for each operation to travel from its origin (remote) region to the target (local) region.  The latency is computed as T2 - T1, where:  T1 is the timestamp when an operation is performed in the remote region, and  T2 is the timestamp when the agent persisted the replicated operation to the local region.
	For each remote region, the latency statistics are reported as the average, minimum, and maximum latency for all the operations from that region.
	If this statistic information is not available, -1 is printed as its output value.

#### Check the persistence of remote data

Table A-3 Output Statistics 2

Statistic	Description
puts	Number of write operations received.
dels	Number of delete operations received.
streamBytes	Total bytes replicated from a remote region.
persistStreamBytes	Reports the total number of bytes that are successfully committed in the local region. This is different from the total bytes replicated from a remote region because in case of any conflicts with operations from other regions, some of the operations may not persist if they fail the built-in conflict resolution rule.
winPuts	Number of write operations performed successfully. More specifically, this statistic excludes the writes that failed to win the conflict resolution rule, in case of a conflict with writes in other regions.
winDels	Number of delete operations performed successfully. More specifically, this statistic excludes the deletes that failed to win the conflict resolution rule, in case of a conflict with deletes in other regions.
incompatibleRows	Number of operations that did not persist because of incompatible table schemas.  This can happen when there is a schema mismatch between the origin region and the region that is trying to replicate the row to its local data store.

#### Monitor the interaction between admin and the agent

Table A-4 Output Statistics 3

Statistic	Description
requests	All the DDL commands executed by the user on an MR table are converted into requests to the agent by the admin. This statistic reports the number of requests posted by the admin.
responses	Number of requests processed and responded by the agent.

#### Monitor multi-region tables

When you execute the show mrtable-agent-statistics command with the -table flag, it returns the table level statistics indicating:

- 1. Persistence of remote data in the local region: This includes the statistics such as puts, dels, winPuts, winDels, streamBytes, persistStreamBytes, and incompatibleRows discussed above.
- 2. Progress of table initialization in each remote region: This is indicated by the state attribute under the Initialization statistics in the output. The table below lists the different possible values for state and their meaning.

Table A-5 Table Initialization States

State	Description
NOT_START	MR table initialization has not started, or there is no need to do initialization. For example, if the agent resumes the stream from an existing checkpoint successfully, there is no need to re-initialize the MR table.
IN_PROGRESS	MR table initialization is ongoing, that is, the MR table initialization has started and the data is being replicated from the remote regions.
COMPLETE	MR table initialization is complete and table transfer is done. The agent is streaming from the remote region.
ERROR	MR table initialization cannot complete because of an irrecoverable error. You can view the error severity in the agent log as WARNING or SEVERE. The agent log can be found in the directory specified in the JSON config file. See Configure XRegion Service.
SHUTDOWN	MR table initialization cannot complete because the service is shut down.

3. Persistence of the table data per remote region:



Table A-6 Output Statistics 4

Statistic	Description
transferStartMs	Timestamp of the initiation of table initialization, in milliseconds.
	If this statistic information is not available, -1 is printed as its output value.
transferCompleteMs	Timestamp of the completion of table initialization, in milliseconds.
	If this statistic information is not available, -1 is printed as its output value.
elapsedMs	The time elapsed from the start of the table initialization until its completion. elapsedMs = transferCompleteMs - transferStartMs
	This statistic is reported in milliseconds. Before the transfer completion, it reports -1 indicating the unavailability of this statistic.
transferBytes	Number of bytes transferred from the remote (origin or source) region to the local (target) region.
transferRows	Number of rows transferred from the remote region to the local region successfully.
expireRows	Number of rows expired before transferring from the remote region. Because of their TTL value, some rows might expire during the replication. Such rows expire by the time they reach the agent. This statistic counts such expired rows.
persistBytes	Reports the total number of bytes that are successfully committed in the local region. This excludes the rows that are not committed in the local region because they failed the built-in conflict resolution rule. In case of row updates, the entire row is counted for this statistic.
persistRows	Reports the total number of rows that are successfully committed in the local region. Similar to the above statistic, the rows that are not committed in the local region because of the built-in conflict resolution rule are excluded for this count.

#### Example

Below are a few examples of the statistics returned by the show mrtable-agent-statistics command with different input parameters.



#### Note:

If any of the statistics information is not available, -1 is reported as the value for that statistic parameter in the output.

#### $\ensuremath{\mathtt{\#}}\xspace$ MR table agent statistics for a specific agent kv-> show mrtable-agent-statistics -agent 0 -json "operation": "show mrtable-agent-statistics", "returnCode": 5000, "description": "Operation ends successfully", "returnValue": { "XRegionService-1 0": { "timestamp": 1592901180001, "statistics": { "agentId": "XRegionService-1 0", "beginMs": 1592901120001, "dels": 1024, "endMs": 1592901180001, "incompatibleRows": 100, "intervalMs": 60000, "localRegion": "slc1", "persistStreamBytes": 524288, "puts": 2048, "regionStat": { "lnd": { "completeWriteOps": 10, "laggingMs": { "avg": 512, "max": 998, "min": 31 }, "lastMessageMs": 1591594977587, "lastModificationMs": 1591594941686, "latencyMs": { "avg": 20, "max": 40, "min": 10 } }, "dub": { "completeWriteOps": 20, "laggingMs": { "avg": 535, "max": 1024, "min": 45 "lastMessageMs": 1591594978254, "lastModificationMs": 1591594956786, "latencyMs": { "avg": 30, "max": 45, "min": 15

```
}
        },
        "requests": 12,
        "responses": 12,
        "streamBytes": 1048576,
        "winDels": 1024,
        "winPuts": 2048
    }
  }
}
# MR table agent statistics for a specific MR table
kv-> show mrtable-agent-statistics -table users -json
  "operation": "show mrtable-agent-statistics",
  "returnCode": 5000,
  "description": "Operation ends successfully",
  "returnValue": {
    "XRegionService-1 0": {
      "tableID": 12,
      "tableName": "users",
      "timestamp": 1592901300001,
      "statistics": {
        "agentId": "XRegionService-1 0",
        "beginMs": 1592901240001,
        "dels": 1000,
        "endMs": 1592901300001,
        "expiredPuts": 200,
        "incompatibleRows": 100,
        "initialization": {
          "lnd": {
            "elapsedMs": 476,
            "expireRows": 100,
            "persistBytes": 6492160,
            "persistRows": 6340,
            "state": "COMPLETE",
            "transferBytes": 8115200,
            "transferCompleteMs": 1592822625333,
            "transferRows": 7925,
            "transferStartMs": 1592822624857
          },
          "dub": {
            "transferStartMs": 0,
            "transferCompleteMs": 0,
            "elapsedMs": -1,
            "transferRows": 0,
            "persistRows": 0,
            "expireRows": 0,
            "transferBytes": 0,
            "persistBytes": 0,
            "state": "NOT START"
          }
```

```
},
    "intervalMs": 60000,
    "localRegion": "fra",
    "persistStreamBytes": 104960000,
    "puts": 100000,
    "streamBytes": 115200000,
    "tableId": 12,
    "tableName": "users",
    "winDels": 745,
    "winPuts": 90000
}
}
}
```

# show namespaces

```
show [AS JSON] namespaces
```

Shows a list of all namespaces in the system.

For example:

```
sql-> show namespaces
namespaces
ns1
  sysdefault
sql-> show as json namespaces
{"namespaces" : ["ns1","sysdefault"]}
```

## show query

```
show query <statement>
```

Displays the query plan for a query.

For example:

```
sql-> show query SELECT * from Users;
RECV([6], 0, 1, 2, 3, 4)
[
   DistributionKind : ALL_PARTITIONS,
   Number of Registers : 7,
   Number of Iterators : 12,
   SFW([6], 0, 1, 2, 3, 4)
   [
    FROM:
    BASE_TABLE([5], 0, 1, 2, 3, 4)
   [Users via primary index] as $$Users
```



```
SELECT:
*
]
```

# show regions

```
show regions statement ::= SHOW [AS JSON] REGIONS
```

The show regions statement provides the list of regions present in a multi-region Oracle NoSQL Database setup. The parameter AS JSON is optional and can be specified if you want the output to be in JSON format.

#### Example 1: Fetching all regions in a multi-region database setup

```
SHOW REGIONS;
regions
   my_region1 (remote, active)
   my region2 (remote, active)
```

#### Example 2: Fetching all regions in a multi-region database setup in JSON format

## show roles

```
show [as json] roles | role <role name>
```

Shows either all the roles currently defined for the store, or the named role.

## show tables

```
show [as json] {tables | table table name}
```

Shows either all tables in the data store, or one specific table, *table\_name*.

Specify a fully-qualified table name as follows:

Entry specification	Description
table_name	Required. Specifies the full table name. Without further qualification, this entry indicates a table created in the default namespace (sysdefault), which you do not have to specify.



Entry specification	Description
parent-table.child-table	Specifies a child table of a parent. Specify the parent table followed by a period (.) before the child name. For example, if the parent table is Users, specify the child table named MailingAddress as Users. MailingAddress.
namespace-name:table-name	Specifies a table created in the non-default namespace. Use the namespace followed by a colon (:). For example, to reference table <code>Users</code> , created in the <code>Sales</code> namespace, enter table_name as <code>Sales:Users</code> .

The following example indicates how to list all tables, or just one table. The empty tableHierarchy field indicates that table t1 was created in the default namespace:

```
sql-> show tables
tables
   SYS$IndexStatsLease
   SYS$PartitionStatsLease
   SYS$SGAttributesTable
   SYS$TableStatsIndex
   SYS$TableStatsPartition
   ns10:t10
   parent
   parent.child
   sg1
   t1

sql-> show table t1
tableHierarchy
   t1
```

To show a table created in a namespace, as shown in the list of all tables, fully-qualify  $table\_name$  as follows. In this case, tableHierarchy field lists namespace ns1 in which table t1 was created. The example also shows how the table is presented as json:

```
sql-> show tables;
tables
   SYS$IndexStatsLease
   SYS$PartitionStatsLease
   SYS$SGAttributesTable
   SYS$TableStatsIndex
   SYS$TableStatsPartition
   ns1:foo
   ns1:t1

sql-> show table ns1:t1;
tableHierarchy(namespace ns1)
   t1
sql-> show as json table ns1:t1;
```



```
{"namespace": "ns1"
"tableHierarchy" : ["t1"]}
```

#### show users

```
show [as json] users | user <user name>
```

Shows either all the users currently existing in the store, or the named user.

## timeout

```
timeout [<timeout ms>]
```

The timeout command configures or displays the request timeout for this session in milliseconds(ms).

The request timeout is the amount of time that the client will wait to get a response to a request that it has sent.

If the optional  $timeout_ms$  attribute is specified, then the request timeout is set to the specified value.

If the optional timeout\_ms attribute is not specified, then the current value of request timeout is displayed.

#### Example A-1 timeout

The following example gets the current value of the request timeout.

```
sql-> timeout
Request timeout used: 5,000ms
```

#### Example A-2 timeout

The following example set the request timeout value to 20000 milliseconds (20 seconds).

```
sql-> timeout 20000
Request timeout used: 20,000ms
```



A shell command may require multiple requests to a server or servers. The timeout applies to each such individual request. A shell command sends out multiple requests and has to wait for each of them to return before the command is finished. As a result, a shell command may have to wait for longer time than the specified timeout and this total wait could be greater than the wait time of the individual request.



#### timer

```
timer [on | off]
```

Turns the measurement and display of execution time for commands on or off. If not specified, it shows the current state of timer. For example:

```
sql-> timer on
sql-> SELECT * from users where id <= 10 ;
+---+
| id | firstname | lastname | age |
+---+
| 8 | Len | Aguirre | 42 |
| 10 | Montana | Maldonado | 40 |
| 6 | Lionel | Church | 30 |
| 3 | Bruno
           Nunez
                    | 49 |
| 2 | Idona
           | Roman | 36 | | |
| 4 | Cooper | Morgan | 39 |
| 7 | Hanae | Chapman | 50 |
| 9 | Julie | Taylor | 38 |
  1 | Dean
            | Morrison | 51 |
| 5 | Troy
           | Stuart | 30 |
+---+
10 rows returned
```

Time: Osec 98ms

## verbose

```
verbose [on | off]
```

Toggles or sets the global verbosity setting. This property can also be set on a per-command basis using the -verbose flag.

## version

version

Display client version information.

