

▶

RUN

📄

SAVE

👤

SHARE

🕒

SCHEDULE

⚙️

MORE

✔️ Query completed.

1

Now, I check to see if the num_of_doors column contains null values using this query:

2

3

4

5

6

7

8

9

10

SELECT

*

FROM

cars.car_info

WHERE

num_of_doors IS NULL;

Press Alt+F1 for Accessibility Options.

Query results

📄 SAVE RESULTS

📊 EXPLORE DATA

↕

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH	PREVIEW		
Row	make	fuel_type	num_of_doors	body_style	drive_wheels	engine_location	wheel_base	length
1	dodge	gas	null	sedan	fwd	front	93.7	157.3
2	mazda	diesel	null	sedan	fwd	front	98.8	177.8

▶

RUN

📄

SAVE

👤

SHARE

🕒

SCHEDULE

⚙️

MORE

✔️ Query completed.

1

I want to check for other potential errors. I used SELECT DISTINCT to check what values exist in a column:

2

3

4

5

6

SELECT

DISTINCT num_of_cylinders

FROM

cars.car_info;

Press Alt+F1 for Accessibility Options.

Query results

📄 SAVE RESULTS

📊 EXPLORE DATA

↕

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH	PREVIEW
Row	num_of_cylinders					
1	four					
2	six					
3	five					
4	three					
5	twelve					
6	two					
7	tow					
8	eight					

RUN

SAVE

SHARE

SCHEDULE

MORE

Query completed.

1

According to the data description, the compression_ratio column values should range from 7 to 23. Let me check:

2

3

4

5

6

7

SELECT

MIN(compression_ratio) AS min_compression_ratio,

MAX(compression_ratio) AS max_compression_ratio

FROM

cars.car_info;

Press Alt+F1 for Accessibility Options.

Query results

SAVE RESULTSEXPLORE DATA

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSEXECUTION GRAPHPREVIEW

Row	min_compression_ratio	max_compression_ratio
1	7.0	70.0

RUN

SAVE

SHARE

SCHEDULE

MORE

Query completed.

1

Run the above query again without the row with 70 to make sure that the rest of the values fall within the expected range of 7 to 23.

2

3

4

5

6

7

8

9

10

11

SELECT

MIN(compression_ratio) AS min_compression_ratio,

MAX(compression_ratio) AS max_compression_ratio

FROM

cars.car_info

WHERE

compression_ratio <> 70;

Press Alt+F1 for Accessibility Options.

Query results

SAVE RESULTSEXPLORE DATA

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSEXECUTION GRAPHPREVIEW

Row	min_compression_ratio	max_compression_ratio
1	7.0	23.0

RUN
 SAVE ▾
 SHARE ▾
 SCHEDULE ▾
 MORE ▾
 Query completed.

```

1  # Check the drive_wheels column for inconsistencies
2
3  SELECT
4    | DISTINCT drive_wheels
5  FROM
6    | cars.car_info;
```

Press Alt+F1 for Accessibility Options.

Query results

[JOB INFORMATION](#)
[RESULTS](#)
[JSON](#)
[EXECUTION DETAILS](#)
[EXECUTION GRAPH](#)
[PREVIEW](#)

[SAVE RESULTS ▾](#)
 [EXPLORE DATA ▾](#)

Row	drive_wheels
1	rwd
2	fwd
3	4wd
4	4wd

RUN
 SAVE
 SHARE
 SCHEDULE
 MORE
 Query completed.

```

1 # It appears that 4wd appears twice in results. However, because I used a SELECT DISTINCT statement to return unique values, this probably means there's an extra space in one of the 4wd entries
   that makes it different from the other 4wd. To check if this is the case, I use a LENGTH statement to determine the length of how long each of these string variables:
2
3 SELECT
4     DISTINCT drive_wheels,
5     LENGTH(drive_wheels) AS string_length
6 FROM
7     cars.car_info;
```

Press Alt+F1 for Accessibility Options.

Query results

SAVE RESULTS
 EXPLORE DATA

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH	PREVIEW															
	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Row</th> <th>drive_wheels</th> <th>string_length</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>rwd</td> <td>3</td> </tr> <tr> <td>2</td> <td>fwd</td> <td>3</td> </tr> <tr> <td>3</td> <td>4wd</td> <td>4</td> </tr> <tr> <td>4</td> <td>4wd</td> <td>3</td> </tr> </tbody> </table>	Row	drive_wheels	string_length	1	rwd	3	2	fwd	3	3	4wd	4	4	4wd	3				
Row	drive_wheels	string_length																		
1	rwd	3																		
2	fwd	3																		
3	4wd	4																		
4	4wd	3																		