Documentation → PostgreSQL 13

Supported Versions: <u>Current</u> (<u>13</u>) / <u>12</u> / <u>11</u> / <u>10</u> / <u>9.6</u> / <u>9.5</u>

Development Versions: devel

Unsupported versions: 9.4 / 9.3 / 9.2 / 9.1 / 9.0 / 8.4 / 8.3 /

8.2 / 8.1 / 8.0 / 7.4 / 7.3 / 7.2 / 7.1

Search the documentation for...

Q

8.3. Character Types

Chapter 8. Data Types <u>Home</u> <u>Next</u>

8.3. Character Types

<u>Up</u>

Table 8.4. Character Types

Prev

| Name | Description |
|----------------------------------|----------------------------|
| character varying(n), varchar(n) | variable-length with limit |
| character(n), char(n) | fixed-length, blank padded |
| text | variable unlimited length |

Table 8.4 shows the general-purpose character types available in PostgreSQL.

SQL defines two primary character types: character varying(n) and character(n), where n is a positive integer. Both of these types can store strings up to n characters (not bytes) in length. An attempt to store a longer string into a column of these types will result in an error, unless the excess characters are all spaces, in which case the string will be truncated to the maximum length. (This somewhat bizarre exception is required by the SQL standard.) If the string to be stored is shorter than the declared length, values of type character will be space-padded; values of type character varying will simply store the shorter string.

If one explicitly casts a value to $character\ varying(n)$ or character(n), then an over-length value will be truncated to n characters without raising an error. (This too is required by the SQL standard.)

The notations varchar(n) and char(n) are aliases for character varying(n) and character(n), respectively. character without length specifier is equivalent to character(1). If character(1) is used without length specifier, the type accepts strings of any size. The latter is a PostgreSQL extension.

In addition, PostgreSQL provides the text type, which stores strings of any length. Although the type text is not in the SQL standard, several other SQL database management systems have it as well.

Values of type character are physically padded with spaces to the specified width n, and are stored and displayed that way. However, trailing spaces are treated as semantically insignificant and disregarded when comparing two values of type character. In collations where whitespace is significant, this behavior can produce unexpected results; for example SELECT 'a '::CHAR(2) collate $^{\prime\prime}$ C" < E' a\n'::CHAR(2) returns true, even though C locale would consider a space to be greater than a newline. Trailing spaces are removed when converting a character value to one of the other string types. Note that trailing spaces are semantically significant in character varying and text values, and when using pattern matching, that is LIKE and regular expressions.

The storage requirement for a short string (up to 126 bytes) is 1 byte plus the actual string, which includes the space padding in the case of character. Longer strings have 4 bytes of overhead instead of 1. Long strings are compressed by the system automatically, so the physical requirement on disk might be less. Very long values are also stored in background tables so that they do not interfere with rapid access to shorter column values. In any case, the longest possible character string that can be stored is about 1 GB. (The maximum value that will be allowed for n in the data type declaration is less than that. It wouldn't be useful to change this because with multibyte character encodings the number of characters and bytes can be quite different. If you desire to store long strings with no specific upper limit, use text or character varying without a length specifier, rather than making up an arbitrary length limit.)

Tip

There is no performance difference among these three types, apart from increased storage space when using the blank-padded type, and a few extra CPU cycles to check the length when storing into a length-constrained column. While $\operatorname{character}(n)$ has performance advantages in some other database systems, there is no such advantage in PostgreSQL; in fact $\operatorname{character}(n)$ is usually the slowest of the three because of its additional storage costs. In most situations text or $\operatorname{character}$ varying should be used instead.

Refer to <u>Section 4.1.2.1</u> for information about the syntax of string literals, and to <u>Chapter 9</u> for information about available operators and functions. The database character set determines the character set used to store textual values; for more information on character set support, refer to <u>Section 23.3</u>.

```
CREATE TABLE test1 (a character(4));
INSERT INTO test1 VALUES ('ok');
SELECT a, char_length(a) FROM test1; -- (1)
  a char_length
 ok
CREATE TABLE test2 (b varchar(5));
INSERT INTO test2 VALUES ('ok');
INSERT INTO test2 VALUES ('good
INSERT INTO test2 VALUES ('too long');
{\tt ERROR:} \quad {\tt value \ too \ long \ for \ type \ character \ varying(5)}
{\tt INSERT\ INTO\ test2\ VALUES\ ('too\ long'::varchar(5));\ --\ explicit\ truncation}
SELECT b, char_length(b) FROM test2;
   b | char_length
                   2
 good
                   5
 too l
```

(1) The char_length function is discussed in Section 9.4.

There are two other fixed-length character types in PostgreSQL, shown in Table 8.5. The name type exists only for the storage of identifiers in the internal system catalogs and is not intended for use by the general user. Its length is currently defined as 64 bytes (63 usable characters plus terminator) but should be referenced using the constant NAMEDATALEN in C source code. The length is set at compile time (and is therefore adjustable for special uses); the default maximum length might change in a future release. The type "char" (note the quotes) is different from char (1) in that it only uses one byte of storage. It is internally used in the system catalogs as a simplistic enumeration type.

Table 8.5. Special Character Types

| Name | Storage Size | Description |
|--------|--------------|--------------------------------|
| "char" | 1 byte | single-byte internal type |
| name | 64 bytes | internal type for object names |

| <u>Prev</u> | <u>Up</u> | <u>Next</u> |
|---------------------|-------------|------------------------|
| 8.2. Monetary Types | <u>Home</u> | 8.4. Binary Data Types |

Submit correction

If you see anything in the documentation that is not correct, does not match your experience with the particular feature or requires further clarification, please use **this form** to report a documentation issue.





<u>Privacy Policy</u> | <u>Code of Conduct</u> | <u>About PostgreSQL</u> | <u>Contact</u> Copyright © 1996-2020 The PostgreSQL Global Development Group