

Difference between BYTE and CHAR in Oracle column datatypes

اگر کاراکترهای پایگاه داده به صورت utf-8 باشند (که استاندارد هم همین هست)، بعضی از کاراکترها بیشتر از 1 بایت حافظه را اشغال می کنند.

اگر دیتاتایپ یک ستون را به صورت VARCHAR2(10 BYTE) تعریف کنیم، به این معنی است که اوراکل می تواند 10 بایت برای ذخیره استفاده کند ولی این 10 بایت به معنی 10 کاراکتر نیست! زیرا بعضی از کاراکترها مثل کاراکترهای غیر انگلیسی، بیشتر از 1 بایت فضا می گیرند .

حال اگر دیتاتایپ یک ستون را به صورت VARCHAR2(10 CHAR) تعریف کنیم، به اوراکل این اجازه را می دهیم که هرچقدر فضا برای ذخیره 10 کاراکتری که بهش دادیم لازم داشت، استفاده کند و دیتا را ذخیره کند و اهمیتی ندارد که این فضا چند بایت است.

پی نوشت 1: فضای Unicode به چند روش می تواند ذخیره شود که utf-8 یکی از آن هاست. تعداد byte ها در این روش متغیر است و بسته به نیاز و کد، کاراکتر می تواند بین 1 تا 4 بایت تغییر کند.

پی نوشت 2: تابع LENGTH در اوراکل، تعداد کاراکترها را نمایش می دهد. مثلاً با کوئری زیر می توانیم بیشترین تعداد کاراکترهای یک ستون را به دست آوریم:

```
SELECT MAX(LENGTH(column_name)) FROM tbl_name;
```

منابع:

<https://stackoverflow.com/questions/81448/difference-between-byte-and-char-in-column-datatypes>

<https://stackoverflow.com/questions/2850807/how-many-bytes-does-oracle-use-when-storing-a-single-character>