Data cleaning with PostgreSQL

Using LEFT, RIGHT and Length functions:

For manipulating the columns' values, we can initially look at three new functions:

1. **LEFT**
2. **RIGHT**
3. **LENGTH**
4. **POSITION**
5. **STRPOS**
6. **UPPER/LOWER**
7. **CAST or ::**

For the sake of use case, consider we have phone number column in our person_contact table with sample value as (235-234-1243**) [**areacode(3) + phonenumber(7)+ hypens(2)**] =** 12 characters in total

LEFT pulls a specified number of characters for each row in a specified column starting at the beginning (or from the left). As you saw here, you can pull the first three digits of a phone number using **LEFT(phone_number, 3)**.

RIGHT pulls a specified number of characters for each row in a specified column starting at the end (or from the right). As you saw here, you can pull the last eight digits of a phone number using **RIGHT(phone_number, 8)**.

LENGTH provides the number of characters for each row of a specified column. Here, you saw that we could use this to get the length of each phone number as **LENGTH(phone_number)**.

POSITION takes a character and a column and provides the index where that character is for each row. The index of the first position is 1 in SQL. If you come from another programming language, many begin indexing at 0. Here, you saw that you can pull the index of a comma as **POSITION(',' IN city_state)**.

STRPOS provides the same result as **POSITION**, but the syntax for achieving those results is a bit different as shown here: **STRPOS(city_state, ',')**.

Note, both **POSITION** and **STRPOS** are case sensitive, so looking for **A** is different than looking for **a**.

Therefore, if you want to pull an index regardless of the case of a letter, you might want to use **LOWER** or **UPPER** to make all of the characters lower or uppercase.

More readings for supported string operations and other functions in Postgres are as:

PostgreSQL: Documentation: 8.1: String Functions and Operators

PostgreSQL CAST Convert From One Data Type Into Another (postgresqltutorial.com)

Using SQL String Functions to Clean Data | Advanced SQL - Mode