

Group 2: Introduction to SQL

Basic Functionality of SQL

Megan Lerman, Gabriel Cunningham, Pradeep Singh and Alexis
Donaldson

March 7, 2018

Goals

- ▶ Find and work with a basic dataset using SQL
- ▶ Introduce some basic SQL functionality for new users
- ▶ Understand what SQL is doing in the background

Find and Download a Basic Dataset

The data that we will be using is from the Consumer Finance Protection Bureau. Click [here](#) for more information. To download the dataset, scroll down and select "Download all complaint data — CSV"

Select

The SELECT statement is used to select data (columns) from a dataset (tables).

```
SELECT ProductName, State, ...  
FROM ConsumerComplaints;
```

Select Distinct

The `SELECT DISTINCT` statement is used to return distinct values.

```
SELECT DISTINCT ProductName, State, ...  
FROM ConsumerComplaints;
```

Where

The WHERE clause is used to filter data. It is also used to extract only the data that fulfills a specified condition.

```
SELECT ProductName, SubProd, ...  
FROM ConsumerComplaints  
WHERE CA;
```

And, Or, Not

The WHERE clause can be combined with AND, OR, and NOT operators.

```
SELECT State, SubProd, ...  
FROM ConsumerComplaints  
WHERE ID AND 83854 ...;
```

Order By

The ORDER BY keyword is used to order our results set.

```
SELECT ProductName, SubProd, ...  
FROM ConsumerComplaints  
ORDER BY ProductName, SubProd, ... ASC—DESC;
```


Insert Into

The INSERT INTO statement is used to insert new records in a table.

```
INSERT INTO ConsumerComplaints (ProductName, SubProd,  
SubIssue, ...)  
VALUES (Mortgage, "Credit Card");
```

Update

Used to modify existing data

```
UPDATE ConsumerComplaints  
SET Tags = NULL  
WHERE Tags = ""
```

Null Values

A null value is a field where data was not entered

```
SELECT Tags  
FROM ConsumerComplaints  
WHERE Tags is NOT NULL
```

Delete

Delete existing records from the table

```
DELETE FROM ConsumerComplaints  
WHERE Tags is NULL
```

Min and Max Values

Select the minimum or maximum from a quantitative field

Select the minimum days between when the complaint was received and when the complaint was sent to the company

```
SELECT MIN DATEDIFF(year, DateSent4, DateRecieved)  
FROM ConsumerComplaints
```

Select the maximum days between when the complaint was received and when the complaint was sent to the company

```
SELECT MAX DATEDIFF(year, DateSent, DateRecieved)  
FROM ConsumerComplaints
```

Count and Average

Count: returns the number of rows that match the condition

Count the number of complaints that came from Older Americans who live in California

```
SELECT COUNT(State) FROM ConsumerComplaints WHERE  
State IS "CA" AND Tags is "Older American"
```

Average: returns the average of a quantitative field

Find the average number of days between receiving the complaint and sending it to the company.

```
SELECT AVG(DATEDIFF(year, DateSent, DateReceived)) FROM  
ConsumerComplaints
```

Like

LIKE is meant to be used with a WHERE statement. If you are searching for something but do not have all the information, it can give you information similar to what you are looking for.

ConsumerComplaints WHERE State LIKE 'a%' A state that starts with A

Wildcards

Wildcards are the characters used with the LIKE function

The % character signifies that it could be any character. The _ represents one character space

In

IN is used with a WHERE statement to search through multiple values

WHERE State IN (CA, CO, CN)

Between

BETWEEN is used to search within a range. It is used with the WHERE statement.

```
SELECT Zip FROM ConsumerComplaints  
WHERE Zip BETWEEN 48382 AND 48384;
```

Group By

The GROUP BY statement is used to group the given data (columns, rows, table) together on the basis of some condition.

```
SELECT Company, State  
FROM ConsumerComplaints  
GROUP BY State
```

Having

The HAVING statement is similar to GROUP BY and is used to do GROUP BY operation with some aggregate function.

```
SELECT Company, State  
FROM ConsumerComplaints  
GROUP BY State  
HAVING Company = "Bank of America";
```

Insert Into Select

INSERT INTO SELECT statement is used to insert data from one table to other tables.

```
INSERT INTO ConsumerComplaintsBackup  
SELECT * FROM ConsumerComplaints
```

SQL Comments

You can add comments in SQL with `--` (two hyphens).
Comments cannot extend to a new line.

```
SELECT Company, State
-- This is Comment
FROM ConsumerComplaints
GROUP BY State
-- give State having company = "Bank of America"
HAVING Company = "Bank of America";
```