Efficient and Dynamic Queries

* SQL Indexes
  + An index is an invisible structure within a database that helps the SQL program process queries faster
  + Creating an index does not change the table data, it simply creates a new data structure that refers to that table.
  + Think of an SQL index like you would the index of a book. You can look up pages in the book with a book index just like you’d look up values in a table with a SQL index
  + It’s the look-up table that speeds up the querying
  + CREATE INDEX
    - The CREATE INDEX command creates a new index. The formal syntax ofr an index on a table looks like this:

CREATE INDEX index\_name ON table\_name;

* + - The formal syntax for an index on a column in a table looks like this:

CREATE INDEX index\_name

ON table\_name (column\_name);

* + - You can also create a UNIQUE INDEX that doesn’t allow duplicate values:

CREATE UNIQUE INDEX index\_name

ON table\_name (column1, column2, ...);

* + A few things to note:
    - If your company is frequently updating a table or performing insert operations, you may want to avoid including an index. These operations take longer on tables with indexes as the index needs to be updated, too.
    - It’s best practice to create indexes only on columns that will be frequently searched against.
    - Indexes should not be used on small tables.
    - Indexes should not be used on columns that contain a high number of null values.
  + Limits of Excel and SQL
    - Excel lets you work with a maximum of 1,048,576 rows. However, as your data set grows, even if it isn’t reaching that limit, your computer may start to slow down. Some of this is related to the specifications of the computer you’re working on, but operations in Excel will start to become tedious on most machines. Functions will take longer to calculate, PivotTables may take longer to create, and charts will take time to render.
    - You could try to perform some of the same functions or groupings in SQL, but SQL programs don’t visualize data, and you’ll eventually reach a point when it just becomes too much for the program to handle.
* Big Data
  + What exactly do we mean by “big data”? As you can see from the previous example, it’s not just thousands of rows in Excel. Usually, big data means millions, or billions of rows of data across several tables and databases. Big data usually takes up terabytes of computer space. Companies like Amazon, Google, Facebook, LinkedIn, and Netflix have this level of data because they are collecting a massive amount of information about each product, user, purchase, video watched, etc. for every single person on their platform. It can be difficult to even imagine that much data!