Class of 01/28/2019

Terminology

- **dataset** - a collection of related data files
  - all files produced by the same source (e.g. U.S Census)
  - in BigQuery, a container of tables (essentially a 1-1 relation with the above definition in our case)
- **relation/entity type** - analogous to "table"
- **row/entity/tuple/record** - analogous terms referring to an entry in a table
- **field/attribute/column** - analogous terms referring to properties of records in a table
- **cell/value** - analogous terms referring to a value of a field for one entity
- **data type** - the type of a certain field to be (e.g. INTEGER, TIMESTAMP)
- **constraint** - a restraint given to a field that forces it to be certain values (e.g. NOT NULL does not allow null values)
  - **primary key** - a constraint that forces all values in the field to be unique and not null
  - **foreign key** - a constraint that forces all values to exist as a value of another field in a table
    - serves as "referencing" one table from another, forming a relationship between two tables
    - parent table has primary key, child table references primary key of parent with a foreign key
- **schema** - the definition of the table (the layout of the table)
- **database** - a collection of tables that have related information

Relationships

- **One to Many (1:m)** one entity in a table can be referenced by multiple entities in another table
  - e.g. A mother can have multiple children, but a child can only have one mother. Mother is said to have a "one to many" relationship with child.
- **One to One (1:1)** one entity in a table can be referenced by only one entity in another table
  - e.g. A husband can only have one wife, and a wife can only have one husband. Husband and wife are said to have a "one to one" relationship.
    - In this table, a parent table should be chosen by the creator
- **Many to Many (m:n)** many entities in a table can be referenced by many entities in another table
○ e.g A student can have many professors, and a professor can have many students
○ junction table - a table representing a many-to-many relationship between two tables (since one field in a table cannot represent many values)
  ■ in a junction table, one field represents a foreign key to one table, and another field the other table. Both fields are primary keys and foreign keys.