Andrew Tran (ant111)

CS 1555 Homework 3

* CUSTOMERS = (SSN, fname, lname, cell\_pn, home\_pn, street, city, zip, state, free\_min, DOB, free\_SMS)
* RECORDS (from\_pn, to\_pn, start timestamp, duration, type)
* STATEMENTS (cell\_pn, start\_date, end\_date, total\_minutes, total\_SMS, amount\_due)
* PAYMENTS (cell\_pn, paid\_on, amount\_paid)
* DIRECTORY (pn, fname, lname, street, city, zip, state)

1. Assuming that the relations CUSTOMERS and PAYMENTS have 10 and 17 tuples, respectively, find the arity and cardinality of the following relations: (For those whose accurate values can not be determined, give the min and max values)
   1. πcell\_pn(Customers):

**Arity = 1; Cardinality = 10**

* 1. πcell\_pn(Payments):

**Arity = 1; Cardinality = 17**

* 1. Customers \* Payments:

**Arity = 14; Cardinalitymin = 0, Cardinalitymax  = 17**

* 1. σpaid on=’2019−09−01’(Payments) ⋈ Payments.cell\_pn=Customers.cell\_pn Customers:

**Arity = 15; Cardinalitymin = 0, Cardinalitymax = 17**

1. Optimize the following relational algebra expression to be more efficient. Please explain your answers.
   1. Πtotal\_minutes, total\_SMS(σcity=’philadelphia’ ∧ Statements.cell\_pn=Customers.cell\_pn(Statements × Customers))

**Πtotal\_minutes, total\_SMS (σcity = ’philadelphia’ (Customers) \* Statements)**

Explanation: Selecting where city = ‘philadelphia’ from just **Customers** will reduce the size of the resulting relation making the join with statements to be more efficient. Similarly, performing a natural join on **Statements** and the reduced **Customer** relation will result in a relation with a max cardinality equal to the cardinality of **Statements**. The cartesian product would produce a relation with a cardinality equal to the product of the cardinalities of **Customers** and **Statements** which would be larger in comparison.

1. Write the relational algebra expression to answer each of the following queries in nesting notation. You can use any date format: ‘2020-01-30’ or ‘01-30-2020’.
   1. List the first and last names of customers who live in Pittsburgh.

**Πfname, lname (σcity = ‘Pittsburgh’ Customers);**

* 1. Retrieve the phone numbers of customers who made calls to people in Pittsburgh.

**ΠCustomers.cell\_pn (((σtype= ‘call’ Records) ⋈ Records.to\_pn = Directory.pn (σcity = ‘Pittsburgh’ Directory)) ⋈ Records.from\_pn = Customers.cell\_pn Customers)**

* 1. List the SSNs of all customers that have ever paid more than 100 in a single payment, and have ever had an amount due more than 50.

**ΠSSN ((Customers \* (σamount\_paid > 100 Payments)) \* (σamount\_due > 50 Statements))**

1. Write the relational algebra expression to answer each of the following queries in sequence notation:
   1. List only once every pair of cell phone numbers which use the same number of SMS in July 2019. **(Assuming all statements start at the 1st of every month)**

**R1 🡨 σstart\_date = ‘07-01-2019’ Statements**

**R2 🡨 σstart\_date = ‘07-01-2019’ Statements**

**R3 🡨 R1 × R2**

**R4 🡨 σR1.cell\_pn != R2.cell\_pn ^ R1.total\_sms = R2.total\_sms (R3)**

**RSLT 🡨 ΠR1.cell\_pn, R2.cell\_pn (R4)**

* 1. Find the SSNs of all customers who received calls from people in Pennsylvania, where they have at least one call duration more than 20.

**PA\_nums 🡨 σstate = ‘Pennsylvania’ Directory**

**Records\_Dur 🡨 σDuration > 20 Records**

**From\_PA 🡨 PA\_nums ⋈ PA\_nums.pn = Records.from\_pn Records\_Dur**

**To\_Customer 🡨 From\_PA ⋈ From\_PA.to\_pn = Customer.cell\_pn Customer**

**RSLT 🡨 ΠSSN To\_Customer**

* 1. List the SSNs for all customers that live in Pittsburgh city and received calls from New York state, but never made calls to New York state.

**Pittsburghers 🡨 σcity = ‘Pittsburgh’ Customers**

**New\_Yorkers 🡨 σcity = ‘New York’ Directory**

**Calls 🡨 σtype = ‘call’ Records**

**To\_Pitt 🡨 Pittsburghers ⋈ Pittsburghers.cell\_pn = Calls.to\_pn Calls**

**NY\_to\_Pitt 🡨 To\_Pitt ⋈ To\_Pitt.from\_pn = New\_Yorkers.pn New\_Yorkers**

**To\_NY 🡨 New\_Yorkers ⋈ New\_Yorkers.cell\_pn = Calls.to\_pn Calls**

**Pitt\_to\_NY 🡨 To\_NY ⋈ To\_NY.from\_pn = Pittsburghers.cell\_pn New\_Yorkers**

**RSLT 🡨 ΠSSN(NY\_to\_Pitt) - ΠSSN(Pitt\_to\_NY)**