**Instructions**

Step1:

MDS\_Exercise1\_FileA.txt (Inventory):

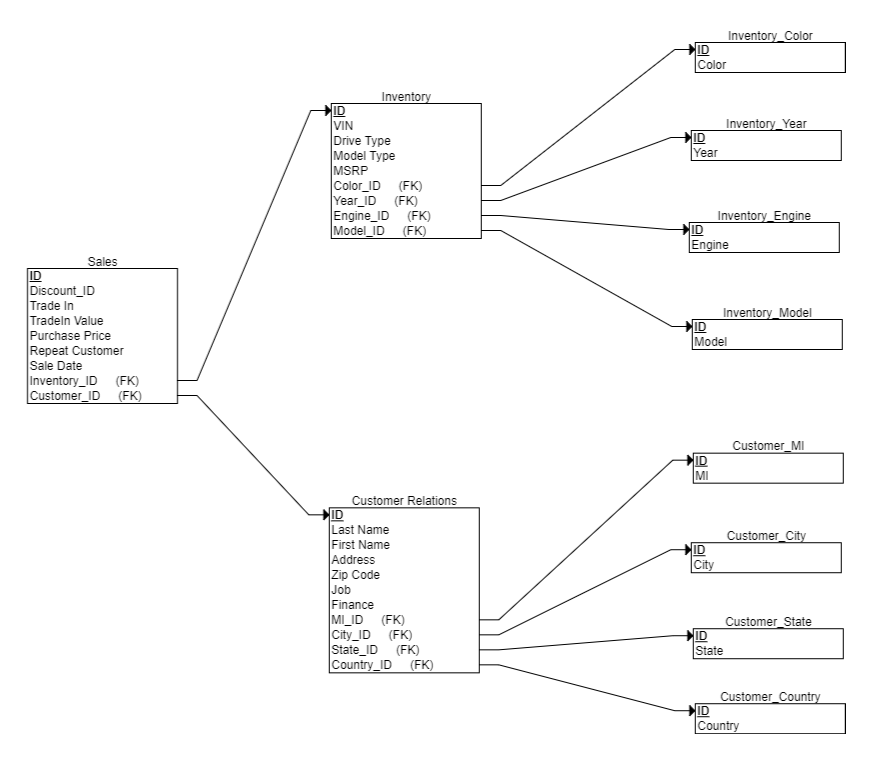
* Information: ID, Model, Model Type, Year, Color, Engine, Drive Type, VIN, MSRP.
* Format and Organization: Text
* Information is Shared to MDS\_Exercise1\_FileB: Model, Model Type, Year, Color, Engine, Drive Type, VIN, MSRP.

MDS\_Exercise1\_FileB.csv (Sales):

* Information: ID, LastName, FirstName, Job, Finance, MI, Address, Zip Code, City, State, Country, SaleDate, Model, Model Type, Year, Color, Engine, Drive Type, VIN, MSRP, Discount, TradeIn, TradeInValue, PurchaseRpice, RepeatCustomer.
* Format and Organization: Comma Separated Values.
* Share is Shared MDS\_Exercise1\_FileA and MDS\_Exercise1\_FileC: Model, Model Type, Year, Color, Engine, Drive Type, VIN, MSRP, LastName, FirstName, Job, Finance, MI, Address, Zip Code, City, State, Country.

MDS\_Exercise1\_FileC.docx (Customer relations):

* Information: LastName, FirstName, Job, Finance, MI, Address, Zip Code, City, State, Country.
* Format and Organization: Word Document
* Information Sharing to MDS\_Exercise1\_FileB: LastName, FirstName, Job, Finance, MI, Address, Zip Code, City, State, Country.



Step4:

* How did you decide to represent the data in the way that you did?

1. First, data cannot be duplicated in each entity (table)
2. Second, the data attributes within an entity should be related
3. Third, decide a thing that can be used to uniquely represent an entity

* Did you leave out any information? If so, why?

Basically, I don’t leave out any information. There are three files and each file only has a few items. I don’t have enough information to decide what information cannot be left out.

* Why did you choose certain things as attributes? As keys?

The reason a thing to be a key is because it can be used to identify a single entity without confusing with other entities. Then others are attributes.

* What were the hardest decisions you had to make in this design process?

The hardest decisions I had to make is to decide what attributes should belong to which tables.

* How does your schema design support data independence?  
  This schema design supports data independence by abstracting out each logical entity into separate tables. The addition or removal of new entities, attributes, or relationships to the conceptual schema should be possible without having to change existing external schemas or having to rewrite existing application programs.
* How may your schema design support the overarching goals of data curation (revisit objectives and activities of Week 1)?
  + Organization
  + Preservation
  + Discoverability
  + Reformatting
  + Sharing

Which curation activities could enhance or sustain the database for future discovery and use for new purposes? What additional activities would you recommend?

Using sample data to test database design

Continuously refining database schema is always necessary because new data or new things will be discovered.