**Assignment 3. Ontologies / ER Diagram Design Exercise**

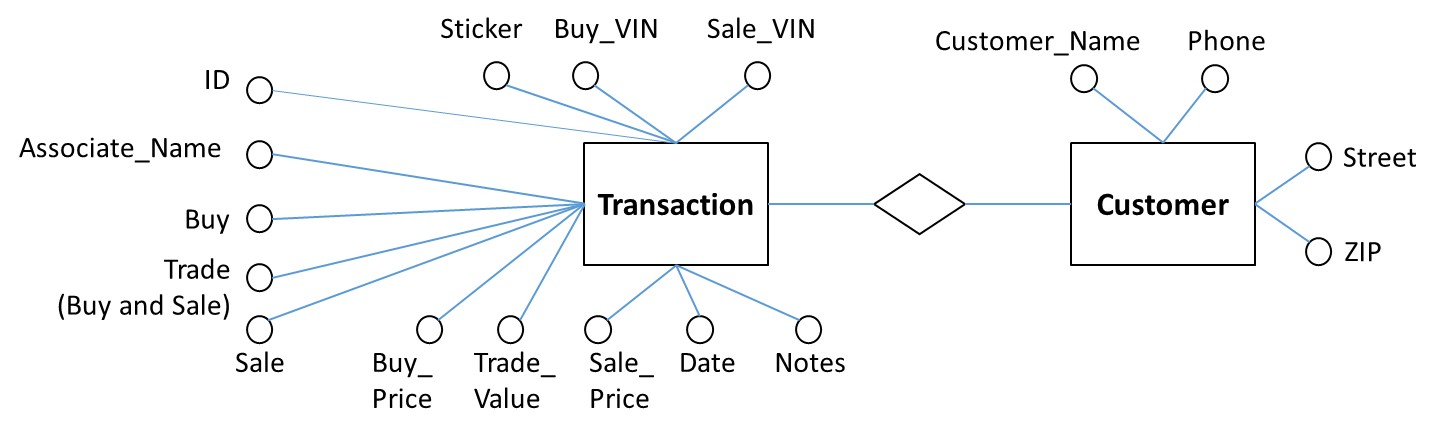
**1. ER Diagram for the Pre-Owned Dealer Database**

The relations in file “*Exercise3\_FaileA.xls*” contain a lot of inconsistencies due to data input errors and need the following data cleaning:

* The BUY, TRADE, and SALE attributes of the TRANSACTIONS relation, when not applicable to a transaction, can have the values of “NULL”, “no”, “-“, or “--” which are variations of NULL handling. In one case an associate even used “y” (transaction 10123465) to denote this fact. For convenience and uniformity, if I had to correct these attribute I would use a Boolean value by placing YES in only one of these three attributes and NO in the rest.
* The BUY\_PRICE, TRADE\_VALUE, STICKER, and SALE\_PRICE attributes of the TRANSACTIONS relation, when not applicable to a transaction, can have the values of “N/A”, “NULL”, or an empty field (also variations of NULL handling). I would use just NULL.
* The BUY\_VIN and SALE\_VIN attributes of the TRANSACTIONS relation, when not applicable to a transaction, can have the values of “”NULL”, “--”, or an empty field (the same type of variations as above). I would use just NULL.
* Sequential transaction numbers do not follow the sequence of dates (e.g. transaction 10123464 occurred later than 10123465). This is quite a mystery to me, and I would enquire the Sales Department about why they do this. Maybe they need to change their procedure and be more consistent.
* Customers’ addresses do not have cities which are more convenient to look up (e.g. online) than zip codes. Need to complete addresses.
* Customer Gandalf does not have a last name. That’s not right. Probably Sales need to call Tolkien’s archivists to clarify that.

Also, it is arguable if ASSOCIATE\_NAME should be an entity or attribute. I considered it an attribute since it is just one column associated with a transaction + there is no such an attribute in the Assignment 1 database. So, even if I make it an entity now, I will still have to convert it to an attribute in the combined schema.

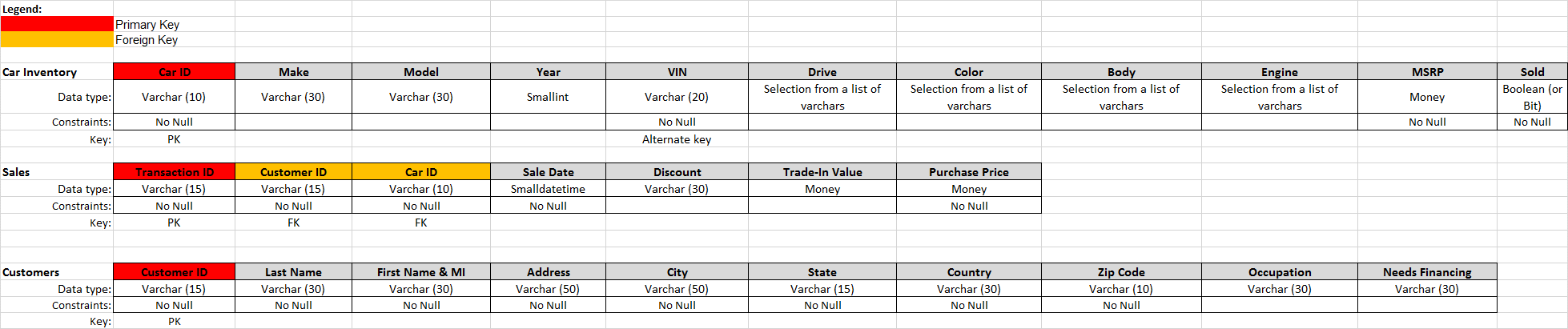
The ER diagram for the pre-owned dealer database looks as follows.



**Figure 1. ER Diagram for Pre-Owned Dealer Database**

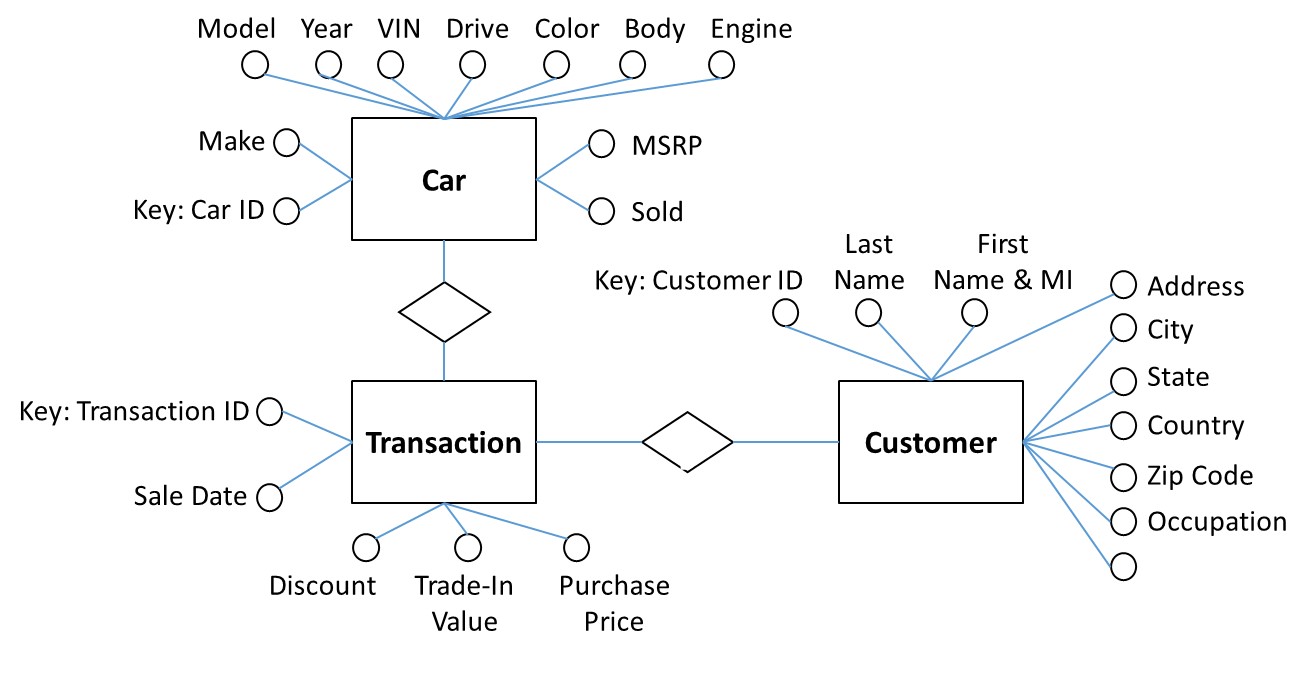
**2. Assignment 1 Schema**

For your convenience, Fig. 2 shows my Assignment 1 schema



**Figure 2. My Assignment 1 Schema**

An ER diagram for this schema is presented in Fig. 3.



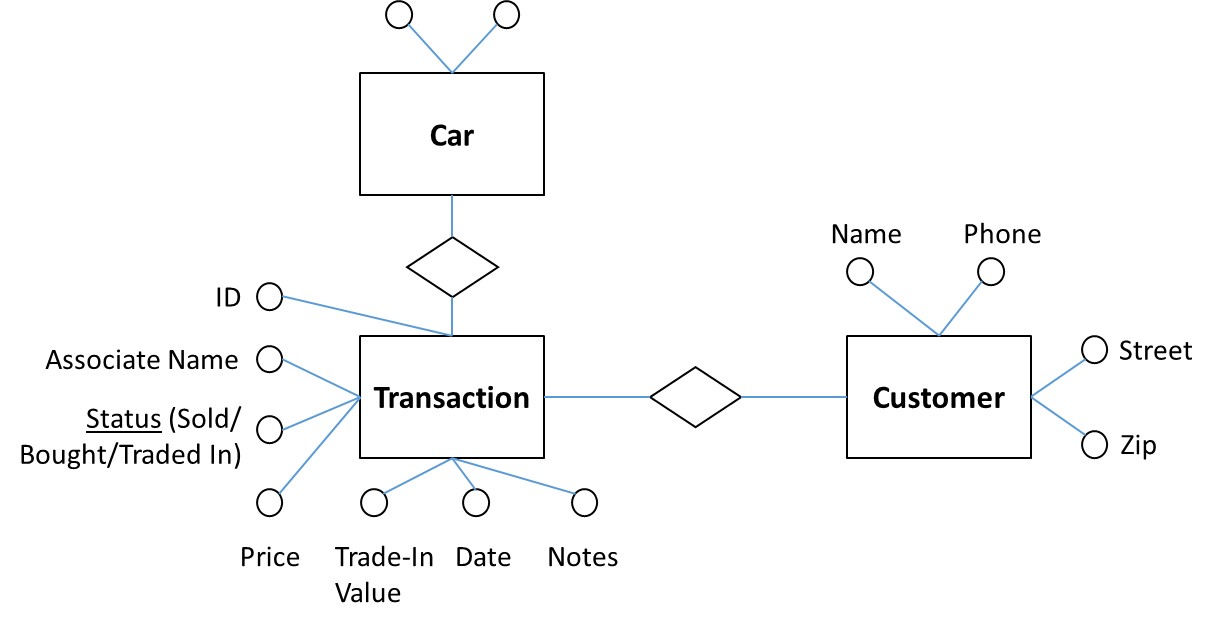
**Figure 3. ER Diagram for My Assignment 1 Schema**

**3. Modification of ER Diagram for Pre-Owned Dealer Database**

I made the following modifications of the initial ER diagram for the pre-owned dealer database:

* Each transaction may be associated with 2 VINs which is confusing. I created a separate entity in the ER diagram called CAR associated with a specific transition. This way, each Car entity will have only one VIN attribute. Each Transaction can be associated with one or more Cars (in case of a sale with trade-in).
* In order to simplify data representation (without loss of information), based on the BUY, TRADE, and SALE attributes from the TRANSACTIONS relation I created just one attribute STATUS in the ER diagram with a list of controlled values: Bought / Sold / Traded In. This can be a more elegant way instead of using three different attributes, the majority of which will have the value of NO, anyway.
* Consequently, the Transaction entity has only one Price attribute corresponding to attribute BUY\_PRICE if the car is purchased, or SALE\_PRICE if it is sold. Attribute Trade In Value of the new Car entity corresponds to attribute TRADE\_VALUE of the initial TRANSACTIONS relation from the pre-owned database.
* The Sticker attribute of CAR has a numeric value if there is a sticker price; otherwise, it is NULL.
* Note: The above inconsistencies and transformations can be used to clean and improve the relations in the relational representation of data shown in Exercise3\_FaileA.xls.

The results of these modifications are shown on Fig. 4



**Figure 4. Step 1: Modified ER Diagram for Pre-Owned Dealer Database**

**4. Combined schema**

You can see now that the two schemas have become quite similar in the structure of entities, and the most significant difference is now the number and composition of attributes for these entities.

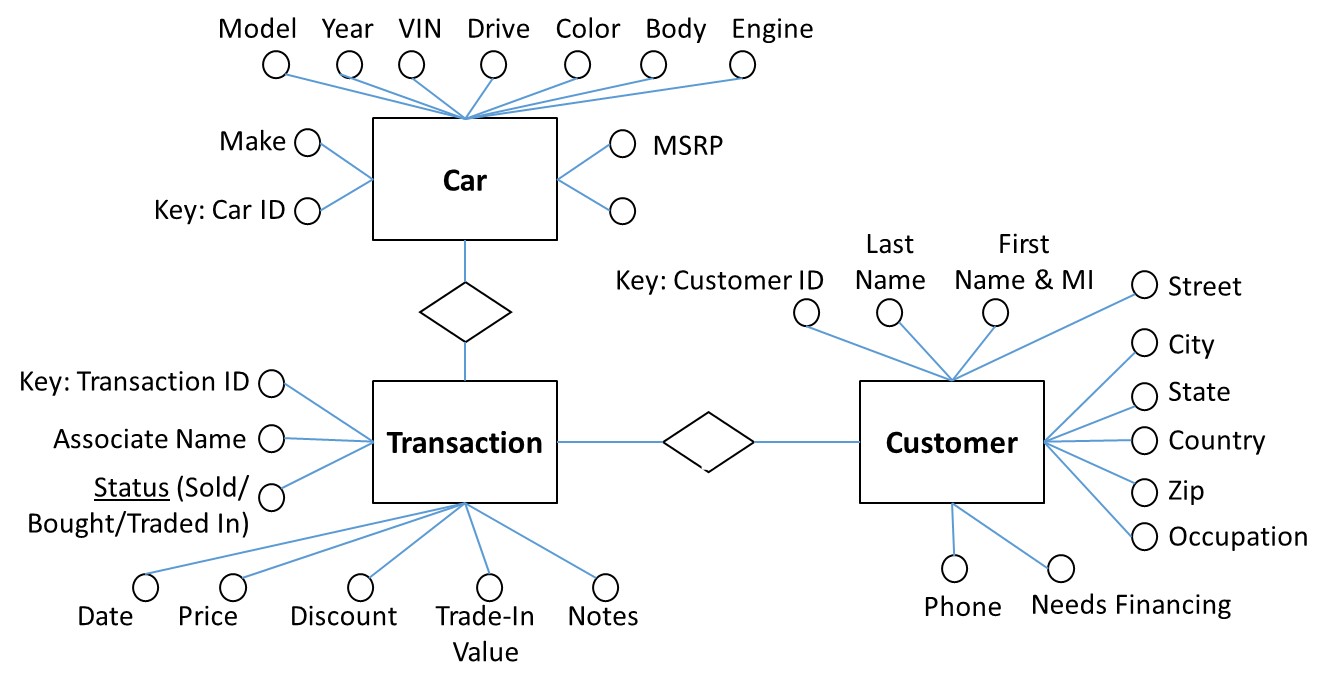
The Sticker and MSRP attributes of the Car entity appear to be synonyms. So are Transaction ID and ID, Sale Date and Date, Trade-In Value attributes of the Transaction entity, and Address and Street, Zip and Zip Code attributes of the Customer entity. In all of these cases I selected the shortest OR the most reasonable name for an attribute in the combined ER diagram.

Since in the Assignment 1 database the cars obtained by the dealer through a trade-in process are not reflected in the database, the Price, Sale Price (as worded in the pre-owned dealer database), and Purchase Price (as worded in Assignment 1) attributes are synonyms in the case of a sale and can be combined. The cars from the Assignment 1 database will simply have no Buy Price, and the Status of the transaction will be Sold.

The Sold attribute of the Assignment 1 Car entity is merged with the Status attribute of the pre-owned Transaction entity.

The Name attribute of the pre-owned database has to be divided into the Last Name and First Name & MI as it was done in the Assignment 1 database. The keys for the Car, Transaction, and Customer entities are Car ID, Transaction ID, and Customer ID, respectively.

Fig. 5 below shows the final combined ER diagram for the two databases based on derivation without any loss of information. There are no homonyms or conceptual overlaps in the two databases.



**Figure 5. Combined ER Diagram**