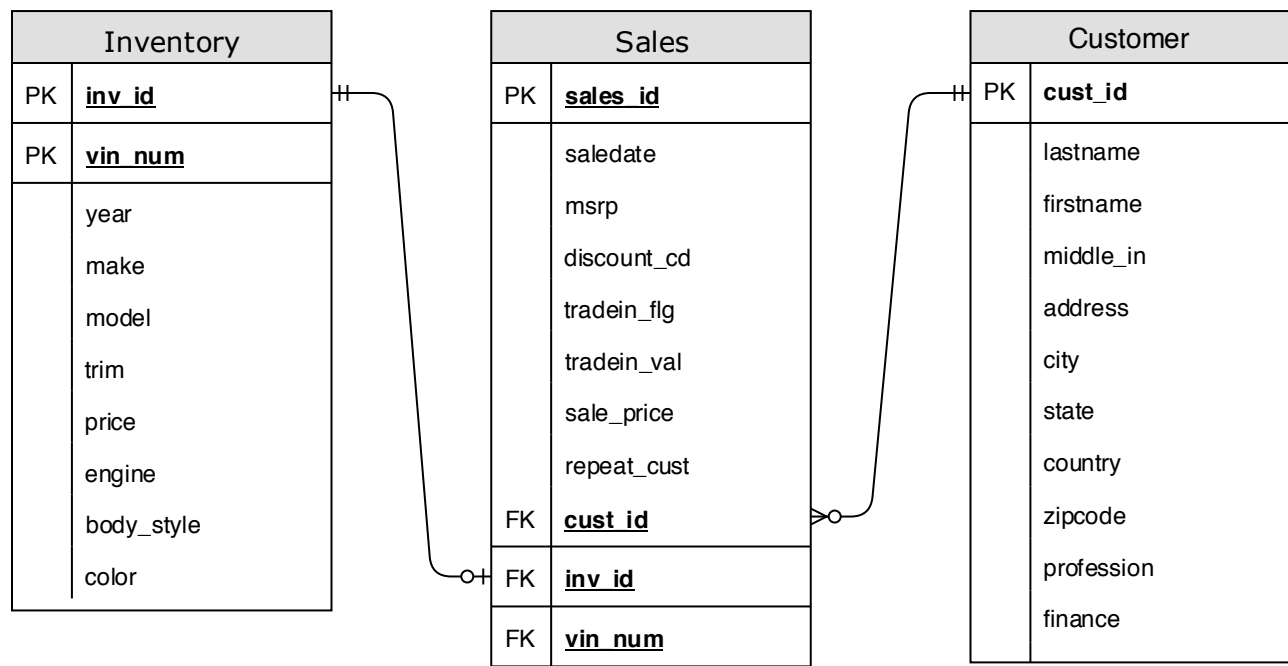# CS598 – Data Curation Assignment 3 – Ved Chugh(vedpc2)

Assignment 3: Ontologies / ER Diagram Design Exercise

As part of the current assignment the scope is to have a Schema integration. So far, we had assignment 1 wherein we had below shared Schema for existing dealer:
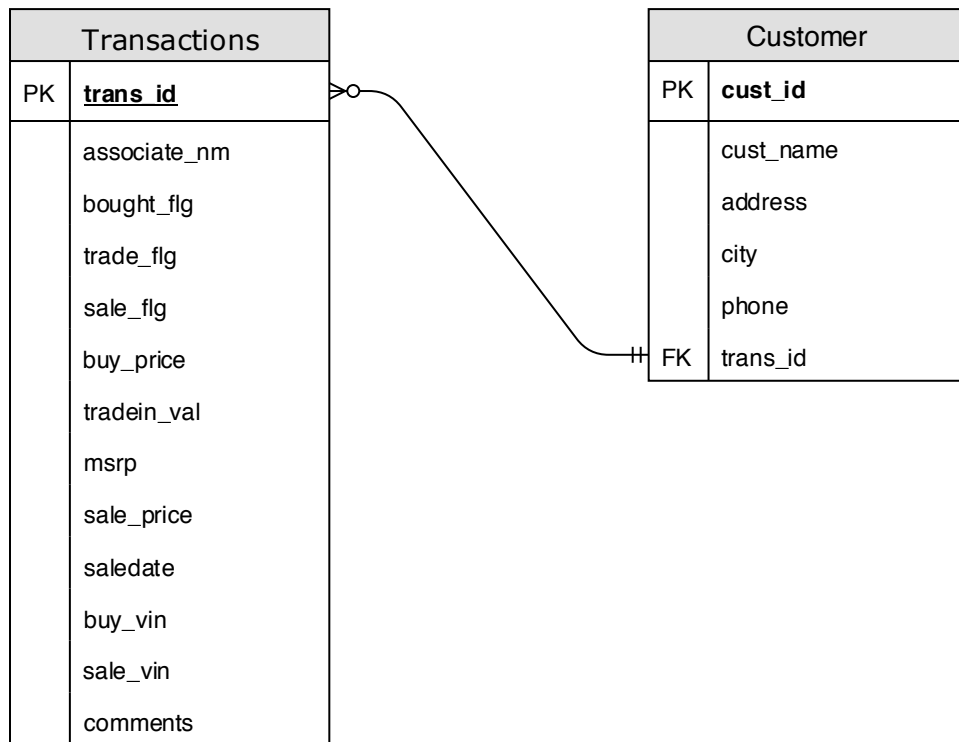
Assignment 1 – Existing Dealer Schema.

Inventory	
PK	**inv_id**
PK	**vin_num**
	year
	make
	model
	trim
	price
	engine
	body_style
	color

Sales	
PK	**sales_id**
	saledate
	msrp
	discount_cd
	tradein_flg
	tradein_val
	sale_price
	repeat_cust
FK	**cust_id**
FK	**inv_id**
FK	**vin_num**

Customer	
PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	city
	state
	country
	zipcode
	profession
	finance

In above schema we have Crow's foot notation used, thus helping us understand below points

a. A sale must belong(minimum) to a customer. Hence "|" added at customer end.
b. A sale can only belong to one customer. Hence "|" added at customer end.
c. A customer doesn't have to participate in sale to exist in database. Hence "o" added at sales end.
d. And further a customer can make many sales. Hence ">" added at sales end.

Similarly, for Sales and Inventory relationship:

a. A sale must have(minimum) an inventory. Hence "|" added at inventory end.
b. A sale can only have to one inventory. Hence "|" added at inventory end.
c. An inventory doesn't have to participate in sale to exist in database. Hence "o" added at sales end.
d. And further an inventory can be part of only one sale record. Hence "|" added at sales end.

Now the Preowned Dealer Schema

Transactions	
PK	**trans_id**
	associate_nm
	bought_flg
	trade_flg
	sale_flg
	buy_price
	tradein_val
	msrp
	sale_price
	saledate
	buy_vin
	sale_vin
	comments

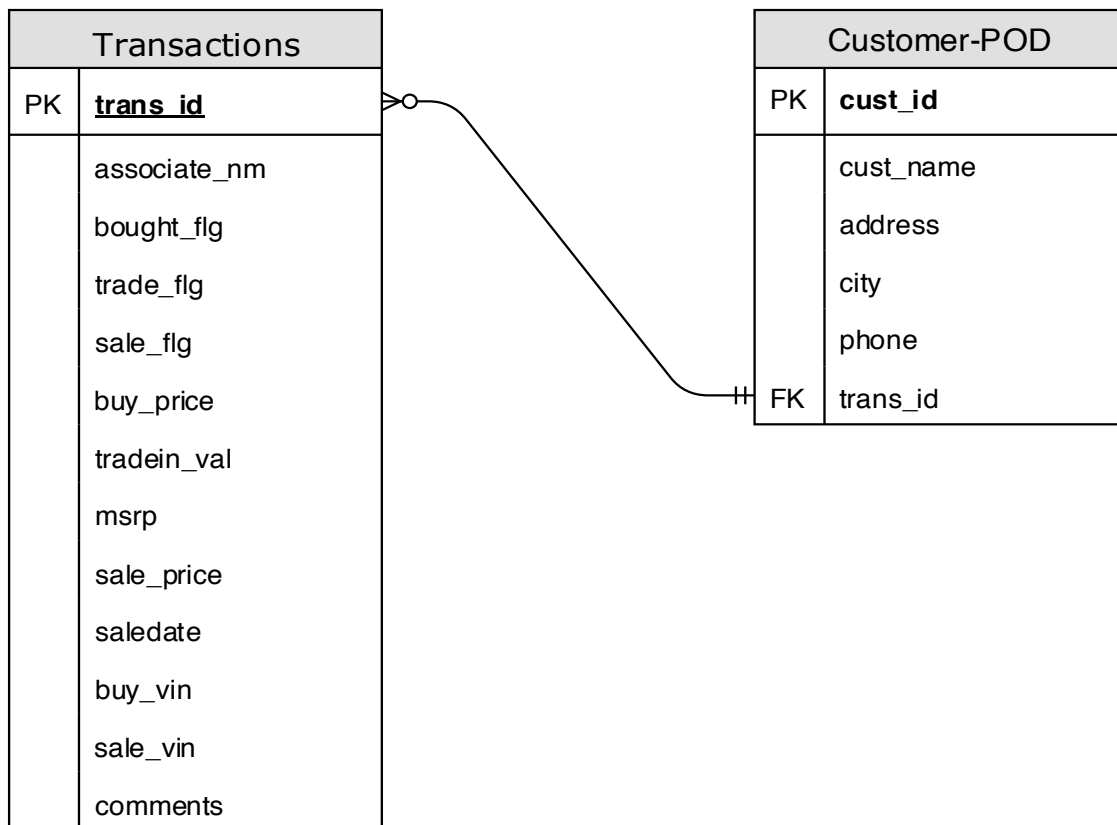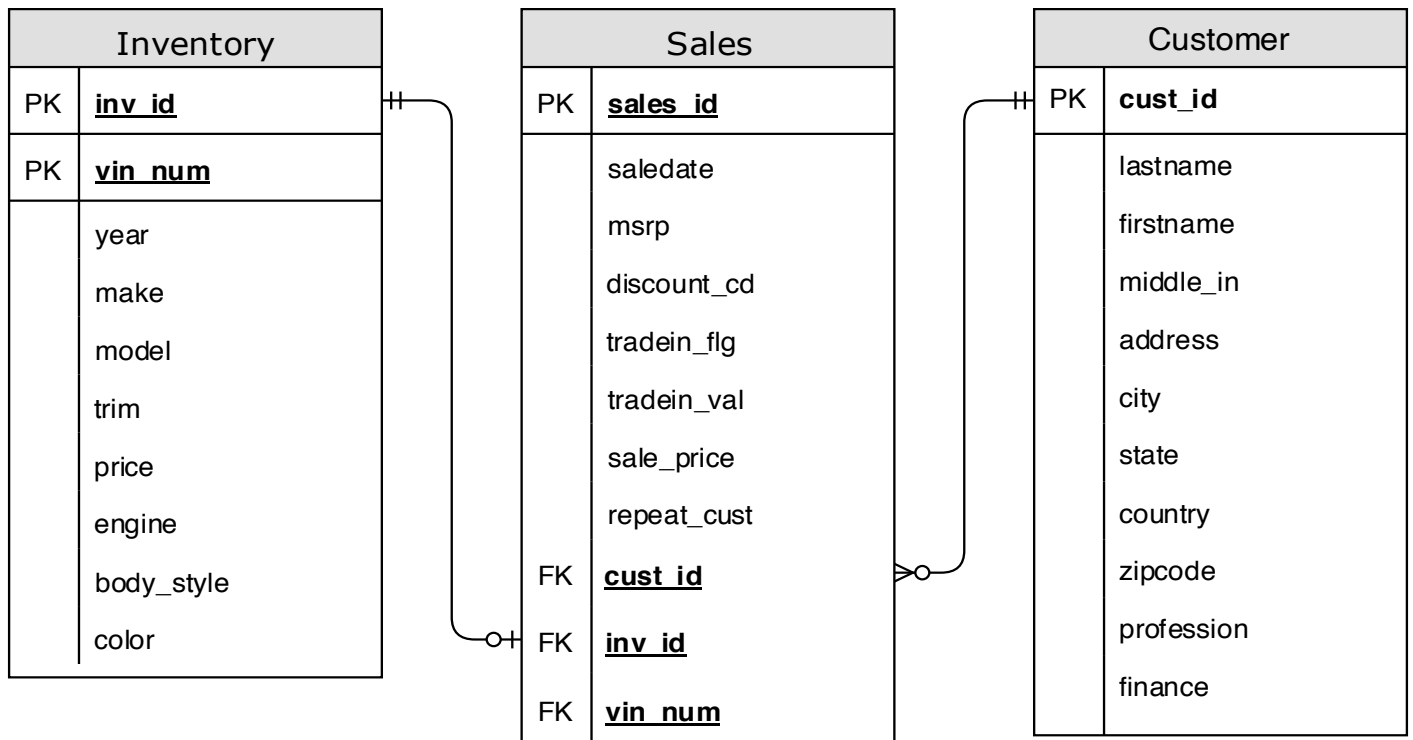| Customer	
PK	**cust_id**
	cust_name
	address
	city
	phone
FK	trans_id

In above schema we have Crow's foot notation used, thus helping us understand below points

a. A transaction must belong(minimum) to a customer. Hence "|" added at customer end.
b. A transaction can only belong to one customer. Hence "|" added at customer end.
c. A customer doesn't have to participate in transaction to exist in database. Hence "o" added at sales end.
d. And further a customer can make many transactions. Hence ">" added at sales end.

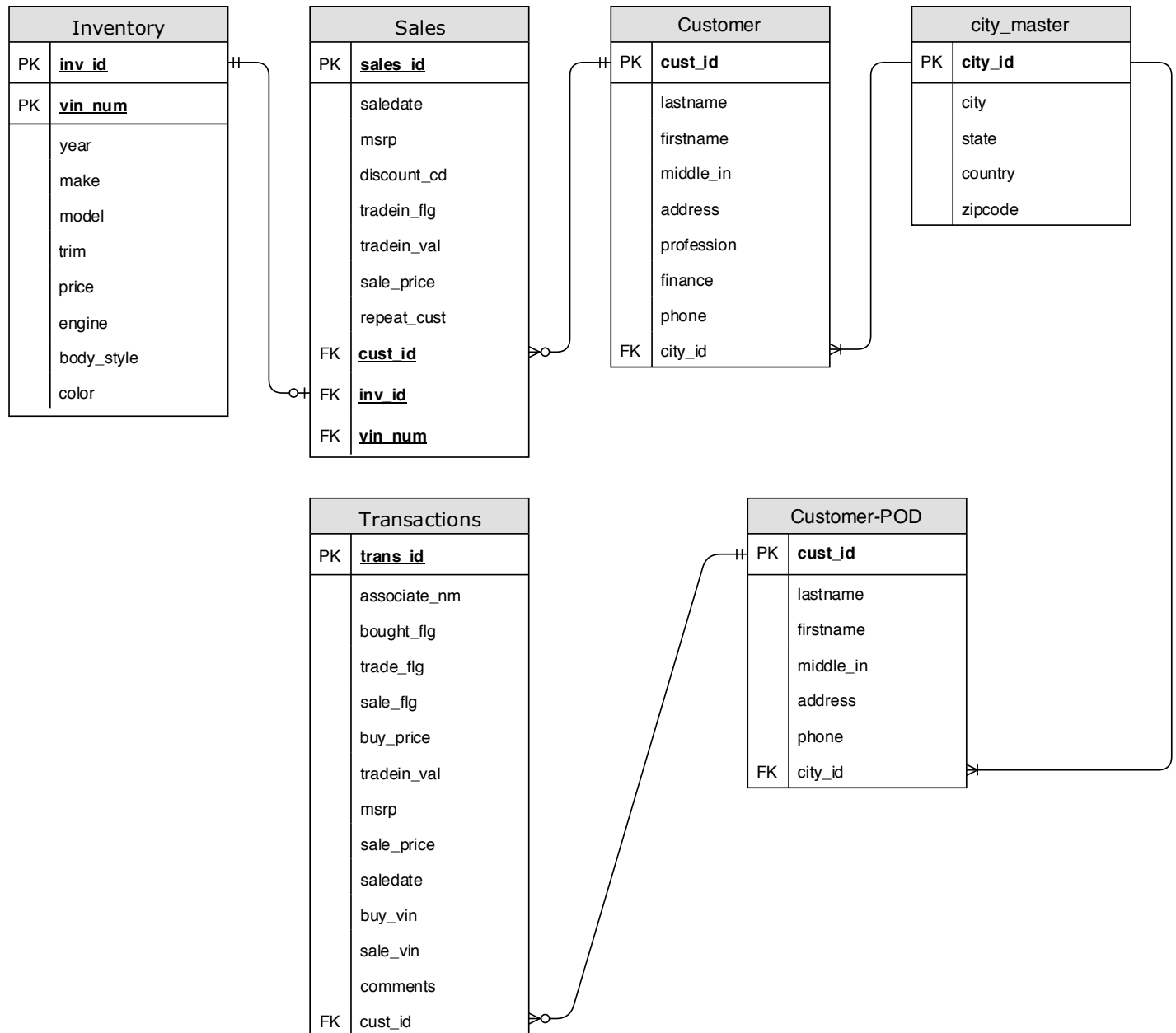Now the task in hand is to have Schema Integration. To start with here are few observations:

a. Both the schema's have Customer table in common.
b. We have attributes which are common.
c. At first glance we see that the Sales table is ⊂ (is a kind of subset) of Transaction table.
d. And further the Inventory details are Entity in our dealer Schema, and attributes of Inventory are in Preowned Dealer Schema.

Inventory

PK	**inv_id**
PK	**vin_num**
	year
	make
	model
	trim
	price
	engine
	body_style
	color

Sales

PK	**sales_id**
	saledate
	msrp
	discount_cd
	tradein_flg
	tradein_val
	sale_price
	repeat_cust
FK	**cust_id**
FK	**inv_id**
FK	**vin_num**

Customer

PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	city
	state
	country
	zipcode
	profession
	finance

Transactions

PK	**trans_id**
	associate_nm
	bought_flg
	trade_flg
	sale_flg
	buy_price
	tradein_val
	msrp
	sale_price
	saledate
	buy_vin
	sale_vin
	comments

Customer-POD

|	
PK	**cust_id**
	cust_name
	address
	city
	phone
FK	trans_id

Step 1: Towards the effort to merge the customer entity, in this step we make city_master as a separate entity containing information about City/State/Country and Zipcode. Thus, below Crow's foot notation:

a. As part of relation a Customer can have only one city associated with them.
b. A city can have multiple customers associated with it. Hence ">" notation is used. Having one to many relationships.

Inventory	
PK	**inv_id**
PK	**vin_num**
	year
	make
	model
	trim
	price
	engine
	body_style
	color

Sales	
PK	**sales_id**
	saledate
	msrp
	discount_cd
	tradein_flg
	tradein_val
	sale_price
	repeat_cust
FK	**cust_id**
FK	**inv_id**
FK	**vin_num**

Customer	
PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	profession
	finance
	phone
FK	city_id

city_master	
PK	**city_id**
	city
	state
	country
	zipcode

Transactions	
PK	**trans_id**
	associate_nm
	bought_flg
	trade_flg
	sale_flg
	buy_price
	tradein_val
	msrp
	sale_price
	saledate
	buy_vin
	sale_vin
	comments
FK	cust_id

Customer-POD	
PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	phone
FK	city_id

Step 2: Towards the effort to merge the customer entity, the above step of creating a city_master further helps us identify that Preowned Dealer Customer is ⊂ of Customer. Thus, we merge the two without loss of information.
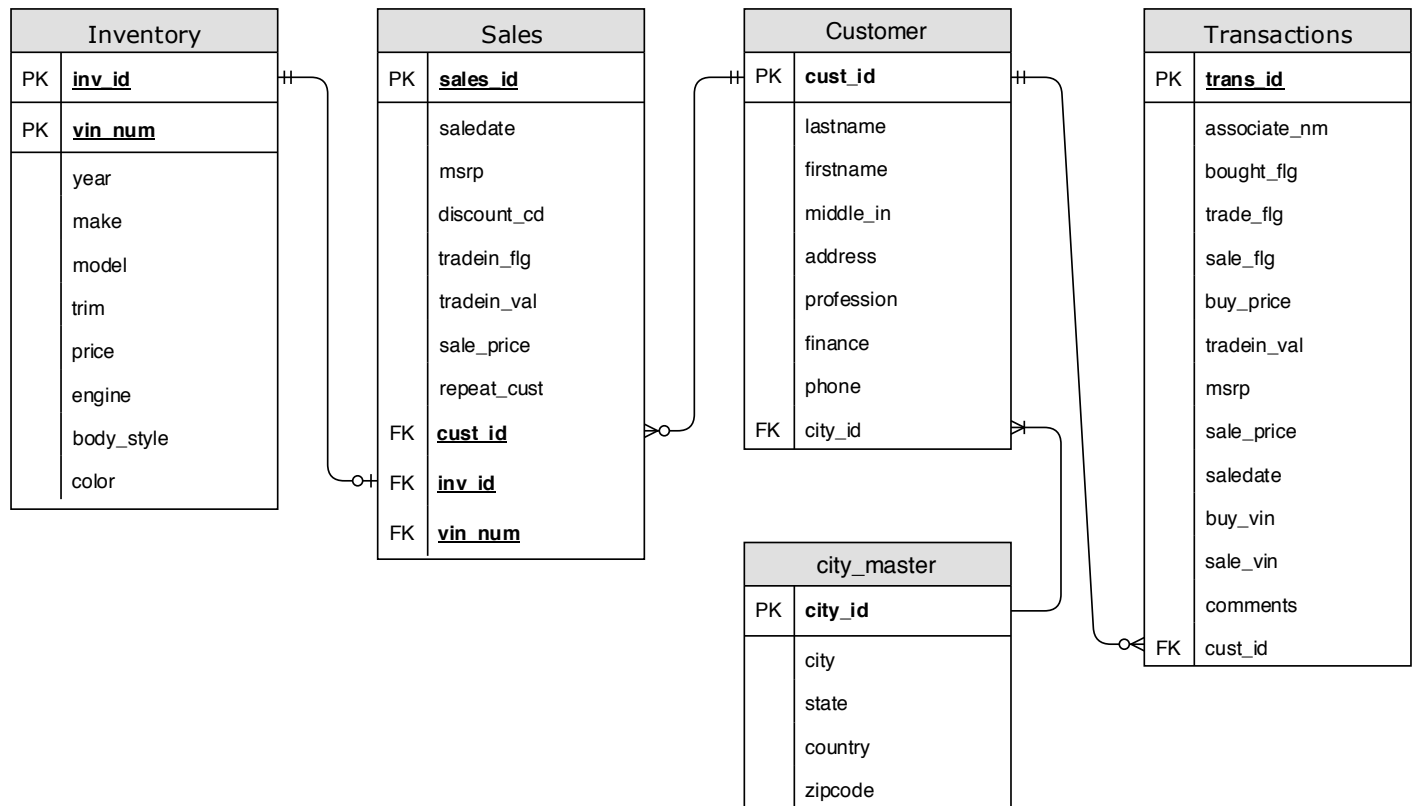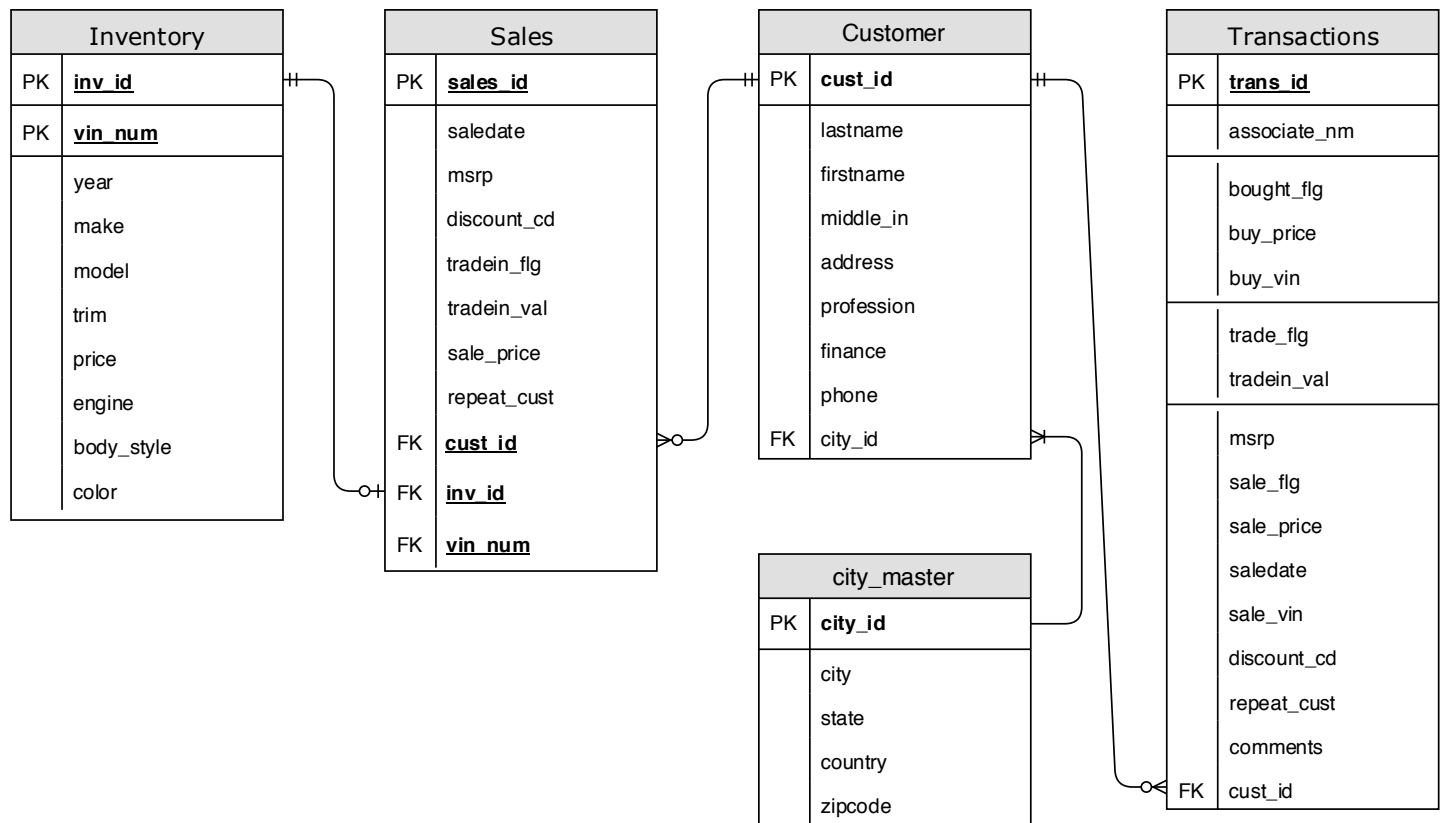

In above schema we have Crow's foot notation used, thus helping us understand below points

a. A transaction must belong(minimum) to a customer. Hence "|" added at customer end.
b. A transaction can only belong to one customer. Hence "|" added at customer end.
c. A customer doesn't have to participate in transaction to exist in database. Hence "o" added at sales end.
d. And further a customer can make many transactions. Hence ">" added at sales end.

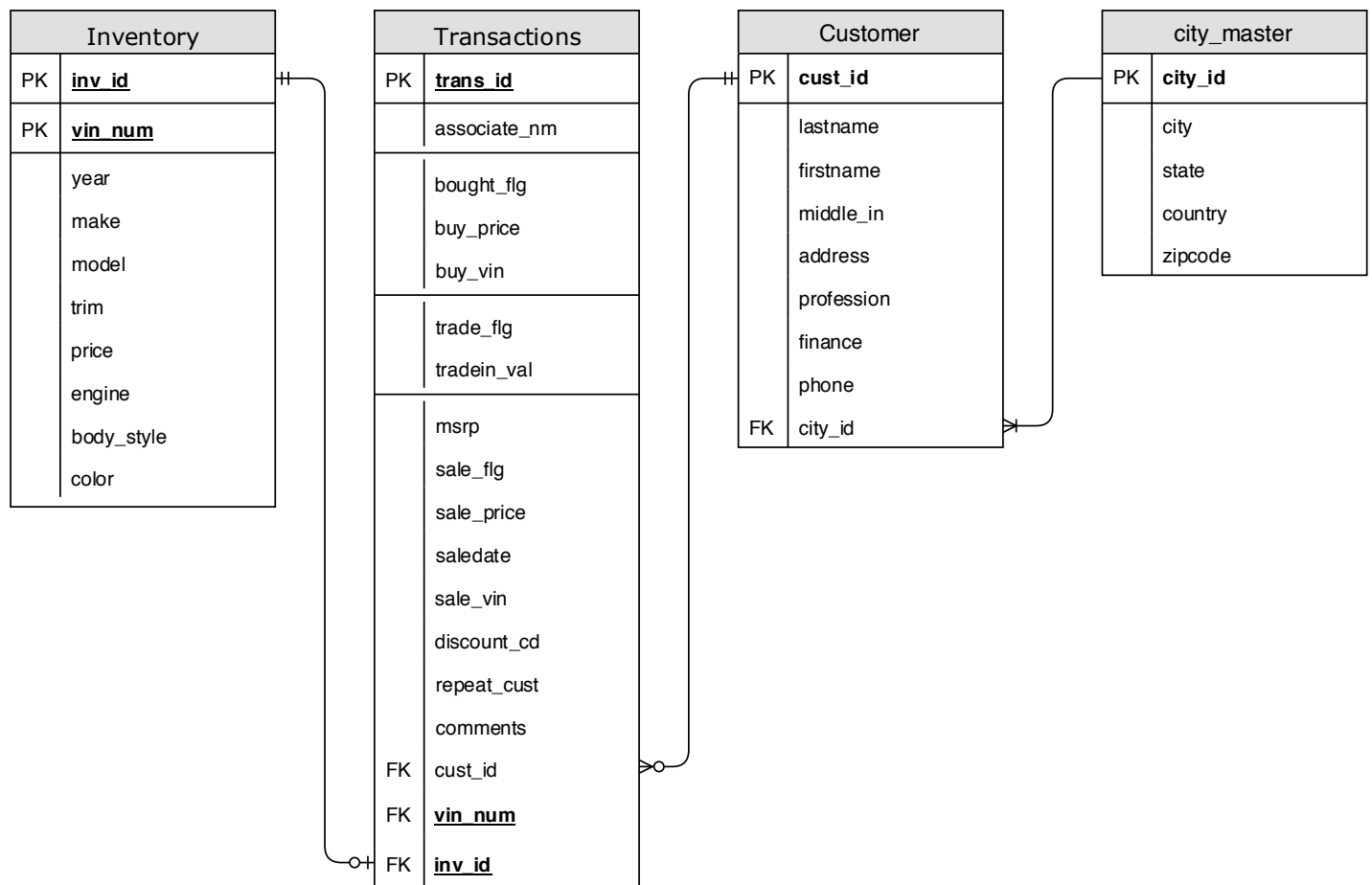We will continue with further with Step 3 on next page.

Step 3: In this step we see with reordering of Transactions table entity that the Sales table is ⊂ (is a kind of subset) of Transaction table.

And therefore, is would be safe to merge the two entities, as all Sales are a kind of transaction and thus merging the two will not lead to any information loss. So, in this step we add extra attributes to Transactions table/entity to help achieve our final merge that will be shown in Step 4.

Inventory	
PK	**inv_id**
PK	**vin_num**
	year
	make
	model
	trim
	price
	engine
	body_style
	color

Sales	
PK	**sales_id**
	saledate
	msrp
	discount_cd
	tradein_flg
	tradein_val
	sale_price
	repeat_cust
FK	**cust_id**
FK	**inv_id**
FK	**vin_num**

Customer	
PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	profession
	finance
	phone
FK	city_id

city_master	
PK	**city_id**
	city
	state
	country
	zipcode

Transactions	
PK	**trans_id**
	associate_nm
	bought_flg
	buy_price
	buy_vin
	trade_flg
	tradein_val
	msrp
	sale_flg
	sale_price
	saledate
	sale_vin
	discount_cd
	repeat_cust
	comments
FK	cust_id

We continue with Step 4 on next page.

6

Step 4: In this we finally combine the schemas for the final Integrated Schema with no information lost.

Inventory	
PK	**inv_id**
PK	**vin_num**
	year
	make
	model
	trim
	price
	engine
	body_style
	color

Transactions	
PK	**trans_id**
	associate_nm
	bought_flg
	buy_price
	buy_vin
	trade_flg
	tradein_val
	msrp
	sale_flg
	sale_price
	saledate
	sale_vin
	discount_cd
	repeat_cust
	comments
FK	cust_id
FK	**vin_num**
FK	**inv_id**

Customer	
PK	**cust_id**
	lastname
	firstname
	middle_in
	address
	profession
	finance
	phone
FK	city_id

city_master	
PK	**city_id**
	city
	state
	country
	zipcode

Below is the link for final merged interactive Schema:

Schema Integration Interactive

7