Assignment 1 – Relation Schema Design Exercise
- RAJA SHEKAR BOLLAM (bollamr2)

[1. 20 Points Narrative Description]

File A
This file has Inventory Information in this case a car. Data on inventory owned by a personal seems to be uniquely identified by an id. This ID is known as VIN as referenced in File B.
It seems to contain information as follows - Year of Purchase of the car, Name of the Manufacturer, Model of the Car, All Wheel Driver or a 4 Wheel Drive, Color of the Car, Number of Car doors, Type of Engine and Price of the Car. Price of the Car is quoted as US Dollars and look consistently formatted in the data set. Data seems to be complete but not arranged in any specific format or Order.
In the context of Completeness, there don't seem to be any missing values or errors
As for the readability, data can be cleaned and arranged in a much better format.
By looking at the data, it is difficult to identify the personal who owns a specific inventory as there is no information related to the personal.

File B
This file seems to have a lot of sales information in regard with the personal and his ownership of a car. It has information related to information on the personal who owns a particular inventory and details on the inventory.
In the context of Completeness, there are several columns missing for some people.
In the context of Redundancy, some of the Inventory details are duplicated from File A.
Since this file is expected to have information related to Sales, data needs to be re-organized.

File C – Personal Information
This file seems to contain Personal Information of the Personal who owns a particular inventory.
Personal Information includes Name with address, Occupation and some other notes.

[2. 10 Points Narrative to design a database schema]

Schema with Tables, Attributes, Data Types, Primary and Foreign Keys.

Table Name	*File_A_Inventory*										
Attributes	**ID**	**VID**	**Year**	**Manufacturer**	**Model**	**WheelDrive**	**Color**	**Doors**	**EngineType**	**MSRP**	
Data Types	INT	VARCHAR(18)	INT	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	INT	VARCHAR(100)	DECIMAL	
Primary Key	*VID*										
Foreign Key											
Table Name	*File_B_Sales*										
Attributes	**LastName**	**FirstName**	**MI**	**SaleDate**	**Discount**	**TradeIn**	**TradeInValue**	**PurchasePrice**	**RepeatCustomer**	***VID***	
Data Types	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	DATE	VARCHAR(100)	BOOLEAN	DECIMAL	DECIMAL	BOOLEAN	VARCHAR(18)	
Primary Key	*Last Name + First Name + MI*										
Foreign Key	*VID*										
Table Name	*File_C_Customer_Relations*										
Attributes	**LastName**	**FirstName**	**MI**	**Address**	**City**	**State**	**Country**	**ZipCode**	**Occupation**	**Notes**	***VID***
Data Types	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(100)	VARCHAR(1000)	VARCHAR(18)
Primary Key	*Last Name + First Name + MI*										
Foreign Key	*VID*										

[3. 15 Points Example of each table]

File A - Inventory

File A	*File_A_Inventory*								
ID	**VID**	**Year**	**Manufacturer**	**Model**	**WheelDrive**	**Color**	**Doors**	**EngineType**	**MSRP**
1	vHxfKmtZ8bSd4JqP5y	2019	Ford	Flex	4WD	Black	4	Internal Combustion	$35,240.00
2	Ab3F3AR5QX4jmxQGNX	2020	Ford	Ecosport S 2.0L	4WD	Red	4	Internal Combustion	$22,080.00
3	S7enznmKTrKsbm4ceC	2019	Tesla	Model S P100 D	AWD	Blue	4	Electric	$133,000.00
4	ZdspCskTUsEMuA5xj4	2017	Tesla	Model S 75D	AWD	Gray	4	Electric	$76,000.00
5	QMsFeqUT38MFLV4NxW	2018	Tesla	Model S 75D	AWD	White	4	Electric	$78,000.00
6	eLqdyxVVA2q5vRZNg5	2018	Tesla	Model S 100D	AWD	White	4	Electric	$96,000.00
7	UW7W4XUcxaMBL2PHqS	2020	Toyota	Corolla Hybrid	FWD	Blue	4	Sedan Hybrid	$23,100.00
8	AQm44N9vhHn6DsWvsr	2019	Toyota	Prius L	FWD	Blue	4	Sedan Hybrid	$23,770.00
9	amdRVQn8AVfrdP48CY	2018	Toyota	Prius	FWD	Silver	4	Sedan Hybrid	$23,475.00
10	3T3zsvzUp5Vm5r2SGm	2018	Toyota	Prius	FWD	Black	5	Hatchback Hybrid	$30,565.00

File B – Sales

File B	*File_B_Sales*								
LastName	**FirstName**	**MI**	**SaleDate**	**Discount**	**TradeIn**	**TradeInValue**	**PurchasePrice**	**RepeatCustomer**	***VID***
Pettigrew	Peter		10/20/2019	EndofYear	Yes	$1,250.00	$17,705.50		Ab3F3AR5QX4jmxQGNX

File C – Customer Relationship

File C	*File_C_Customer_Relations*									
LastName	**FirstName**	**MI**	**Address**	**City**	**State**	**Country**	**ZipCode**	**Occupation**	**Notes**	***VID***
Pettigrew	Peter		55 Shadow Canyon Trl	Indianapolis	IN	USA	46077	Librarian	Needs financing	Ab3F3AR5QX4jmxQGNX

[4. 30 Points Process for creating the database schema and tables. Describe decisions]

Relationship between data files
Upon careful observation on the data provided File A, File B, File C, we come down to the following conclusion on the relationship between these 3 data files.
File B has all the data related to a person's information, Inventory (car) he possess and his purchase information o the inventory. A VID column in File B is an Unique ID that can be used to identify the owner in File A. The First and Last Name in File B can used to get the address of the owner from File C.

Below is the Process followed for creating the database schema. We have divided entire data in to 3 tables. File_A_Inventory, File_B_Sales, File_C_Customer_Relations

File A - Inventory
1. This table needs to have information about the purchased car.
2. VIN is used as Unique Key/ID to identify a record.
3. Year, Manufacturer, Model, WheelDrive, Color, Doors, EngineType, MSRP are identified as the attributes related to a car. Hence they are group in this table.

File B – Sale
1. This table needs to identify a person who needs to buy/bought a car.
2. The inventory details need to brought from File A if required, based on the created VIN number.
3. We need to have some personal details related to the person and some sales related information on the purchased item.
4. We identified SaleData, Discount, TradeIn, TradeInValue, PurchasePrice, RepeatCustomer, VIN as information related to Sale of a Car.
5. VIN Is identified as foreign key that can be used to retrieve car information from File A table.
6. Last Name, First Name, MI are used as a primary key for this table. Same key can be used as a Foreign key to File C to retrieve customer related information.

File C – Customer Relations
1. This table need to have information related to the person, and the kind of customer he is to the company.
2. As a result we have identified Address, City, State, Country, ZipCode, Occupation and Notes as the required attributes for this table.
3. Notes has customer specific information such as "Needs Loan", "Inquiring about Car" etc.
4. Last Name, First Name, MI are used as a primary key for this table. Same key can be used as a Foreign key to File B to retrieve customer related information.

As for the Data Types for each attributes are concerned we have used "VARCHAR" with a length of 100 for all character related entries. INT for number related attributes, DECIMAL/BOOLEAN for items related to this kind of data and DATE data type for SaleDate.