**MIS-300 Enterprise Information Systems/Database**

**Richard Huang**

**Professor: Dr. Nichelle**

**November 18, 2018**

**Final Project Milestone one**

1. **Analysis**

**A. Explain what the data actually is and how the data correlates back to the question you intend to answer.**

These databases are about the Iowa Liquors store’s sales records, and each file represents the sales history of each salesman/saleswomen. These databases include the following information: invoice, date, store number, store name, address, category number, category name, vendor number, vendor name, item number, item name, item description, quantity amount of each pack, cost, retail price, amount sold in bottle, and total sales price. I will be able to retrieve the information from these databases above to find out the total of 2015 sales by month.

**B. Analyze the data and identify any areas that are not aligned properly. Be sure that you adjust these later to ensure that they are matched up properly.**

After reviewing these databases, “Iowa\_liquor\_sales\_52358” has no sales amount. “Iowa\_liquor\_sales\_52216” has been calculated incorrectly by the calculation of state bottle retail multiply the bottles sold. “Iowa\_liquor\_sales\_52747” has incorrect information, whose category code does not match category name.

“Iowa\_liquor\_sales\_52772” date format is different from others.

**C. Describe how you will pull all of the data together in order to make sure it works collectively and will report correctly for your needs.**

Based on my question: “Total of 2015 sales by month”, I would need to collect the date column, sales amount column. I would need to make sure all date formats are consistent and I would need to calculate the sales amount if there was no sales amount by state bottle retail multiply bottle sold.

**D. Based on your analysis, provide a list of potential questions you could answer for management.**

Who was the best salesman/saleswomen of the 2015?

Who was our biggest vendor/supplier?

What was the top sale brand in our store?

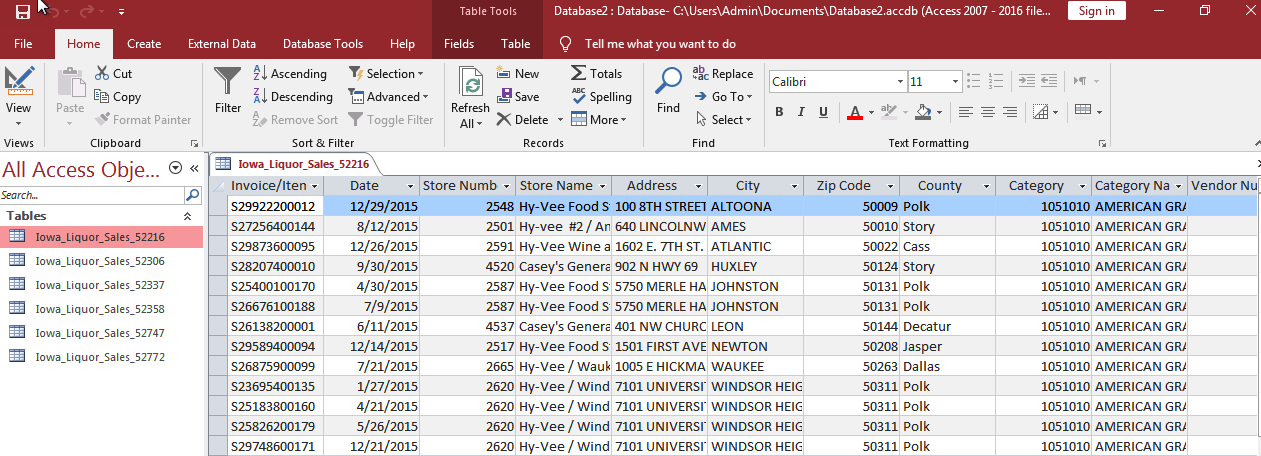
What was the top sale product in December 2015?

What was the best selling product by month?

1. **Compiling and Cleaning the Data**

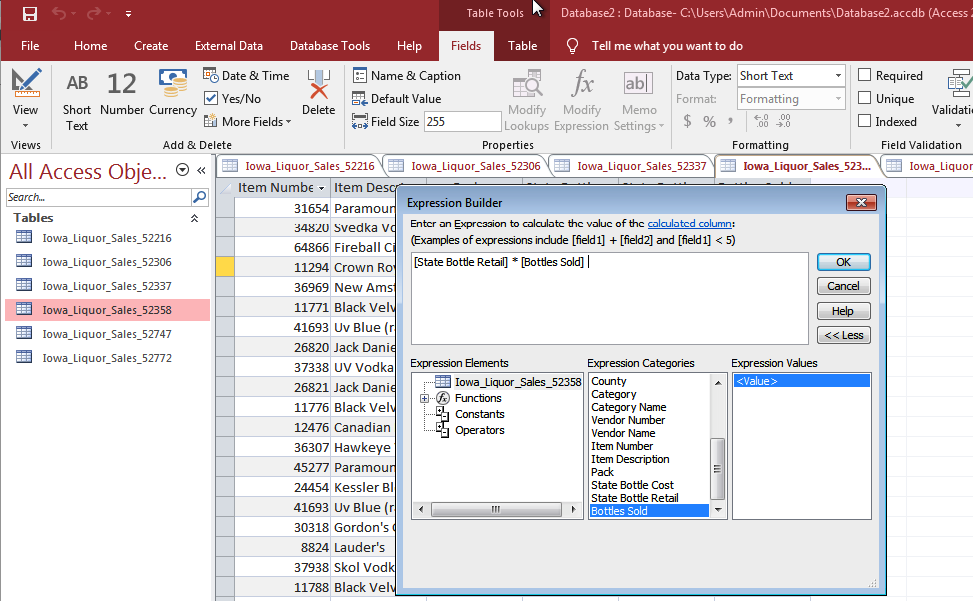
**A. Compile the spreadsheets you have analyzed and load them into Access in order to create a table.**

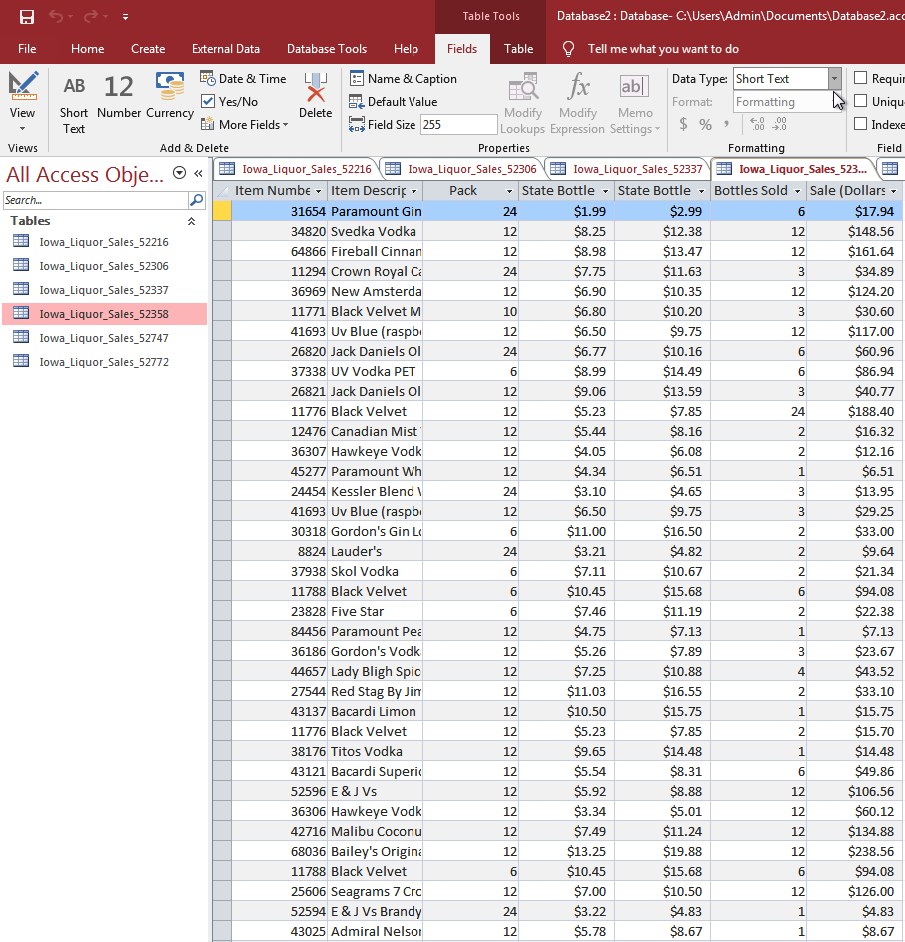
I have loaded all the excel spreadsheets into the access.

****

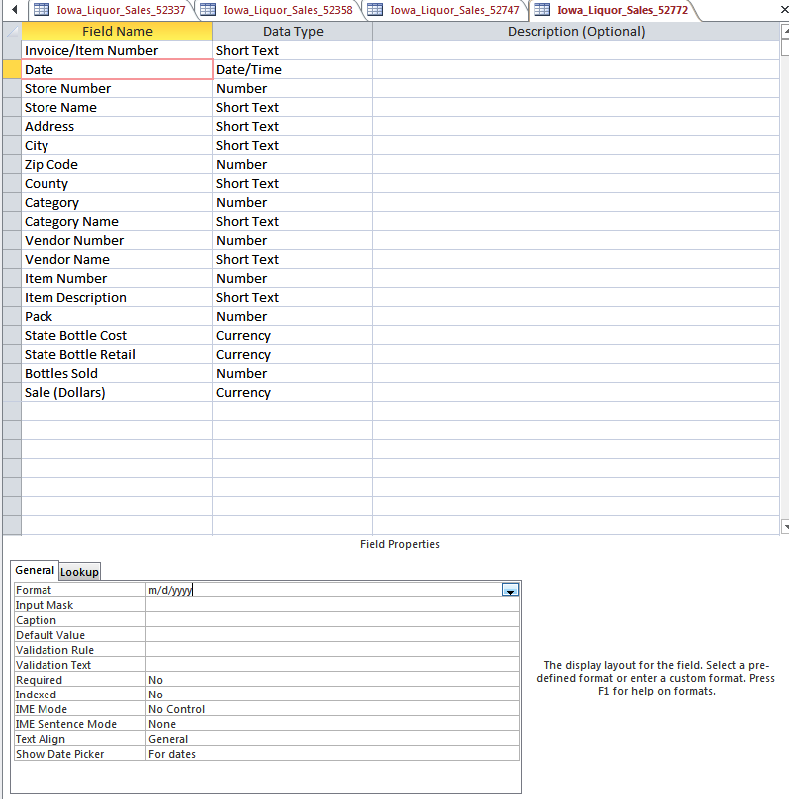
**B. Create a query to maintain the integrity of the original data and clean the data. For example, if you have different names within your table that actually mean the same thing, you will use the query to make the necessary changes.**

1. I would need to calculate the sales amount since there was no sales amount on “Iowa\_Liquor\_Sales\_52358” by inserting calculated field with [State Bottle Retail] \* [Bottle Sold]

****

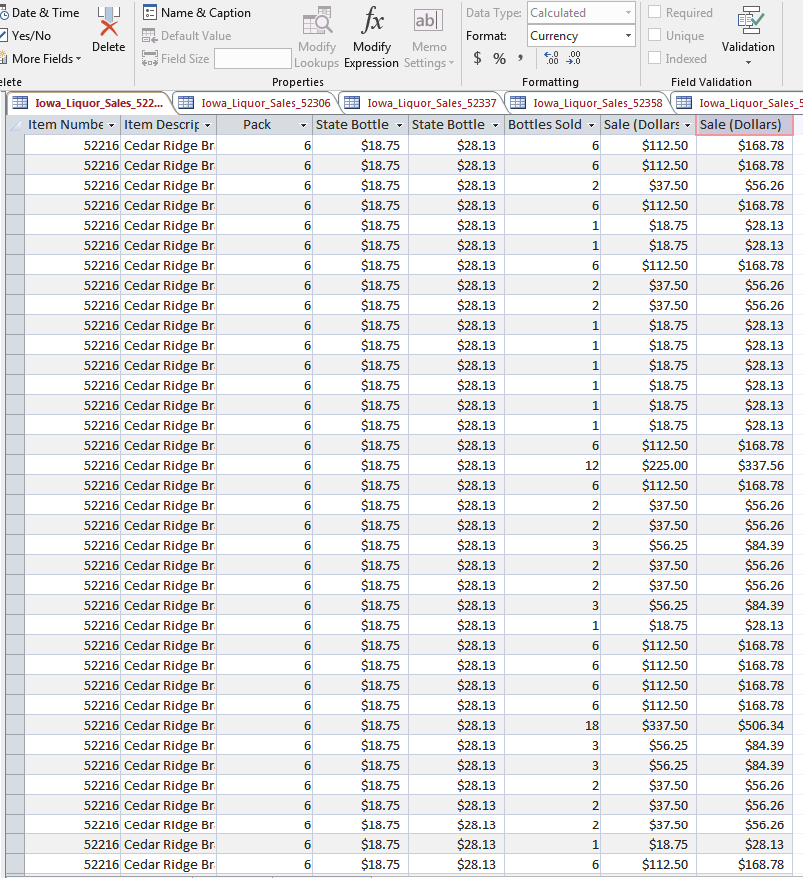
****

2. I would need to correct date format on “Iowa\_Liquor\_Sales\_52772” from “d-mmm-yyyy@” to “ m/d/yyyy” in design view section.

****

**C. While cleaning the data, be sure to look for and fix any incorrect, incomplete, and erroneous data.**

I had use the same method that I used on question B to insert the correct calculation of “Iowa\_Liquor\_Sales\_52216” table. I would also need to delete the incorrect “Sale (Dollars)” field.

****

**D. Based on your initial analysis and your compiled and cleaned data, explain how it will effectively answer the business question.**

After initial analysis and compiling and cleaning data, the data has been correct and consistent. It would greatly increase the compatibility of the data, and it would calculate the accurate data to answer the business questions.