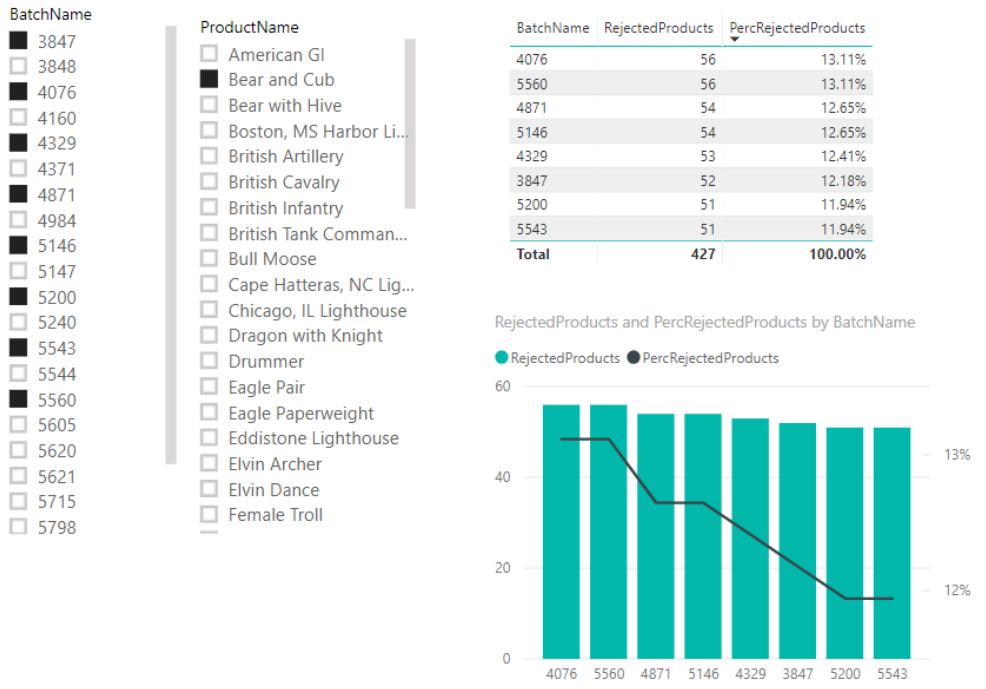
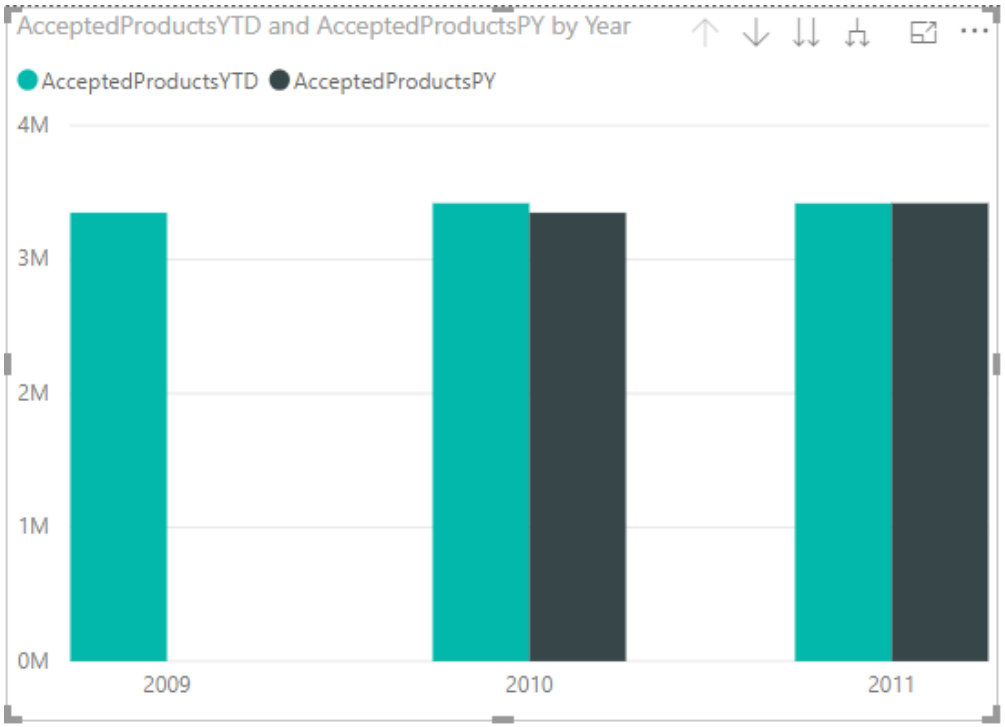
Assignment 2

Part I: DAX function and visual charts

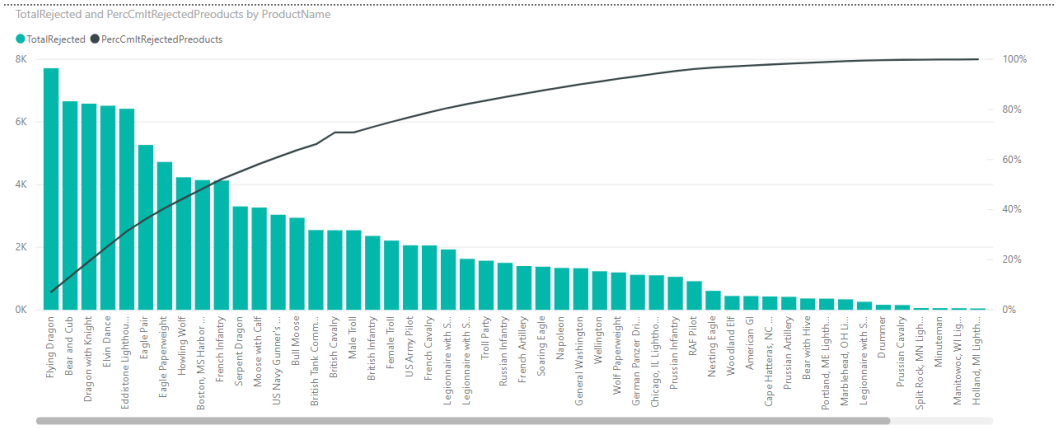
1. Create "Assignment 2-YourLastName-YourFirstName.pbix" in Power BI Desktop
2. Load all the dimension tables and the ManufacturingFact table from your individual database into the project. The relationships between tables may already created after the loading. Examine the relationships on the Model view of the project. If not, create the relationships (refer to the file "Create database tables-final data mart tables.pdf" for the relationships)
3. In DimProduct, create the following columns and measures (Some formulas are provided in a separate file. Try to finish this part without referencing the provided code)
    1) a new column TotalAccepted, which is the total accepted number of the product
    2) a new column TotalRejected, which is the total rejected number of the product (Use CALCULATE function)
    3) a new column TotalProduced, which is the total accepted plus total rejected
    4) a new column RejectionRank, which ranks the products according to the total number of rejection of a product in the ManufacturingFact table
    5) a new column AcceptanceRank, which ranks the products according to the total number of an accepted product in the ManufacturingFact table.
    6) a new measure GrandRejected for rejected number of all products
    7) a new measure for the cumulative number of rejected products: CmltRejectedProducts
    8) a new measure PercCmltRejectedProducts for the cumulative percentage of the number of rejected products
4. In ManufacturingFact, create the following columns and measures
    1) a new measure for the number of rejected products: TotalRejected (the sum of RejectedProducts)
    2) a new measure for 7 day moving average of rejected products: 7DMVRejected. Use ManufacturingFact[DateOfManufacture] in the formula for all date parameters
    3) a new measure for the number of accepted products in the current year: AcceptedProductsYTD
    4) a new measure for the number of accepted products in the prior year: AcceptedProductsPY
    5) a new measure for the percentage of the number of rejected products in one batch compared with the number in selected batch names: PercRejectedProducts
5. Create four report pages.
    1) A page that includes
        a. a slicer of ProductName
        b. a slicer of BatchName, which includes the top 20 batch names in terms of the total rejected products
        c. a matrix table that shows the total rejected number of the selected product in each batch and each batch's PercRejectedProducts in the selected batches
        d. a chart that shows the selected batches, their rejected number of the selected product and the PercRejectedProducts



2) A page that compares AcceptedProductsYTD and AcceptedProductsPY



3) A page that shows the rejected number of each product and the PercCmltRejectedProducts in a combination chart (bar and line). This page can be used to see which products are the cause of 80% total rejected products (a Pareto chart)



4) A page that shows the moving average number of rejected products. The initial view looks like the following.

With the Drill Mode turned on, you can drill down to the day level as shown in the image below

Part II: Data mining

Use the SQL Server Management Studio to examine the tables in the database MIS422Dataset_DM_Assignment (The column descriptions are in the Excel files on Canvas). Challenge yourself with two business questions that data mining techniques can help answer. For example, as a car dealer, what price should I offer for a used car or what factors influence the price of a used car? In Visual Studio, create a data mining project and use at least two learning methods to build models for each question. When you create the data source, make sure on the Impersonation Information tab, select Use the service account option. After the project is created in Visual Studio, right click the project name in the Solution Explorer window, choose Properties. On the Property window, click Deployment under Configuration Properties. Then, in the right pane for the Deployment, enter misbi.cbe.wwu.edu\multi for the Server property and assignment2-YourWindowsLoginName for the Database property.

Write a report in a Word document: "Assignment 2-DM-YourLastName-YourFirstName.docx". In the report, present, but not limited to, the following for each question:

1. The business question that data mining can help to answer
2. The data mining tasks needed for answering the question
3. The learning methods you choose to carry out the task
4. Examine the results from different tabs in each model viewer and present summarized insights from the results
5. Which model you want to use to make the prediction for new data? Why do you choose that model?

Submission: submit the following files to Canvas

1. Assignment 2-YourLastName-YourFirstName.pbix
2. Assignment 2-DM-YourLastName-YourFirstName.docx