Adventure Works Production Data Analytics Project

Data analysis using SQL and PowerBI

Executive Summary

The aim of this analysis is to focus on Adventure Works Cycles, a Microsoft fictitious company which manufactures and sell bicycles in North America, Europe, and Asia. I will analyze the Production department from 2011 to 2014 to gain insights on several aspects including total quantity produced by year, quantity by production unit, orders by product, and production units with highest and lowest production capacity.

Methodology

Before performing the analysis, the data was collected through a public domain, then wrangled to make sure it's cleaned, reliable and error-free. After that, I explored the different variables, proceeded to data visualization to better capture trends and insights and finally draw conclusion.

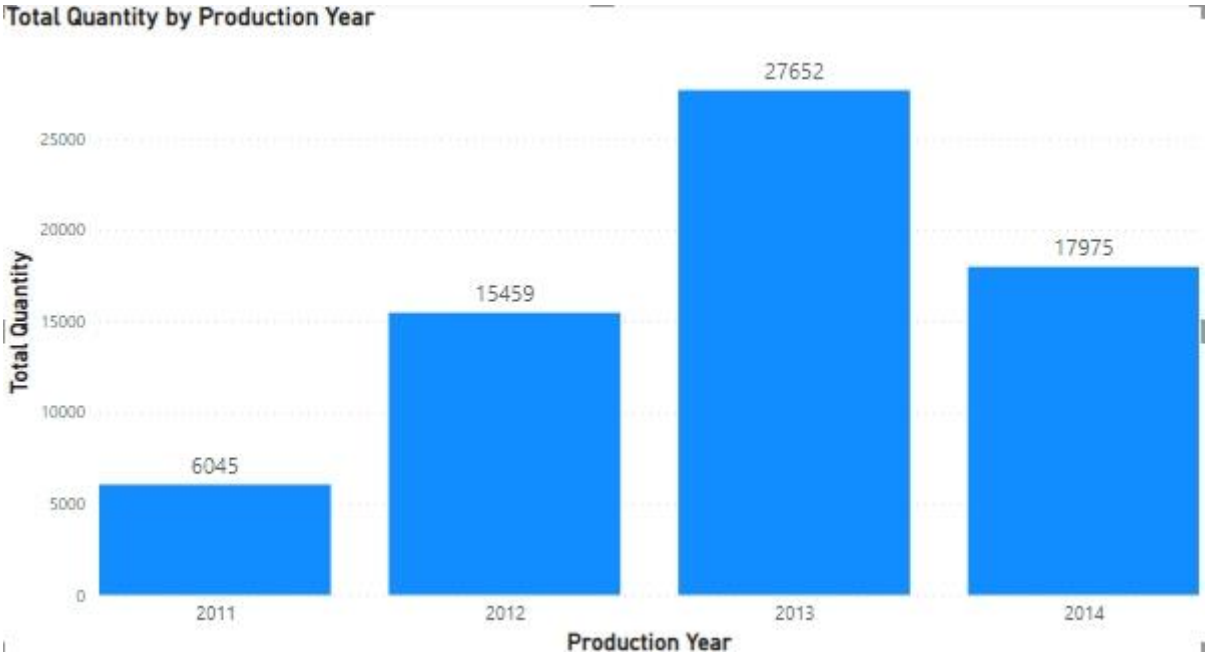
Tools: SQL, PowerBI

Data collection and preparation

The data to explore and analyze was made available through Microsoft website https://learn.microsoft.com/en-us/sql/samples/adventureworks-install-configure?view=sql-server-ver16&tabs=ssms. I proceeded to the collection and the storage of data by making sure they meet the requirements in terms of integrity, reliability, credibility and security.

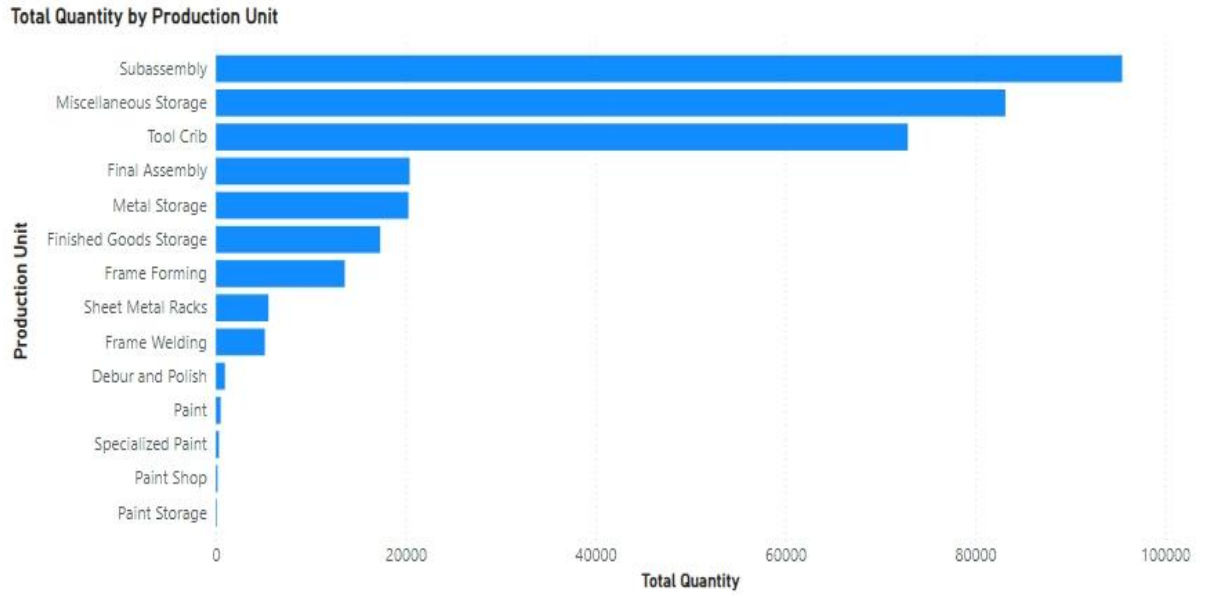
Data Analysis And Visualization

```
--Showing product quantity produced by year

SELECT Year(ActualEndDate) AS Production_Year
,COUNT(Year(ActualEndDate)) AS Total_Quantity
FROM AdventureWorks2014.Production.WorkOrderRouting Work
GROUP BY Year(ActualEndDate)
```



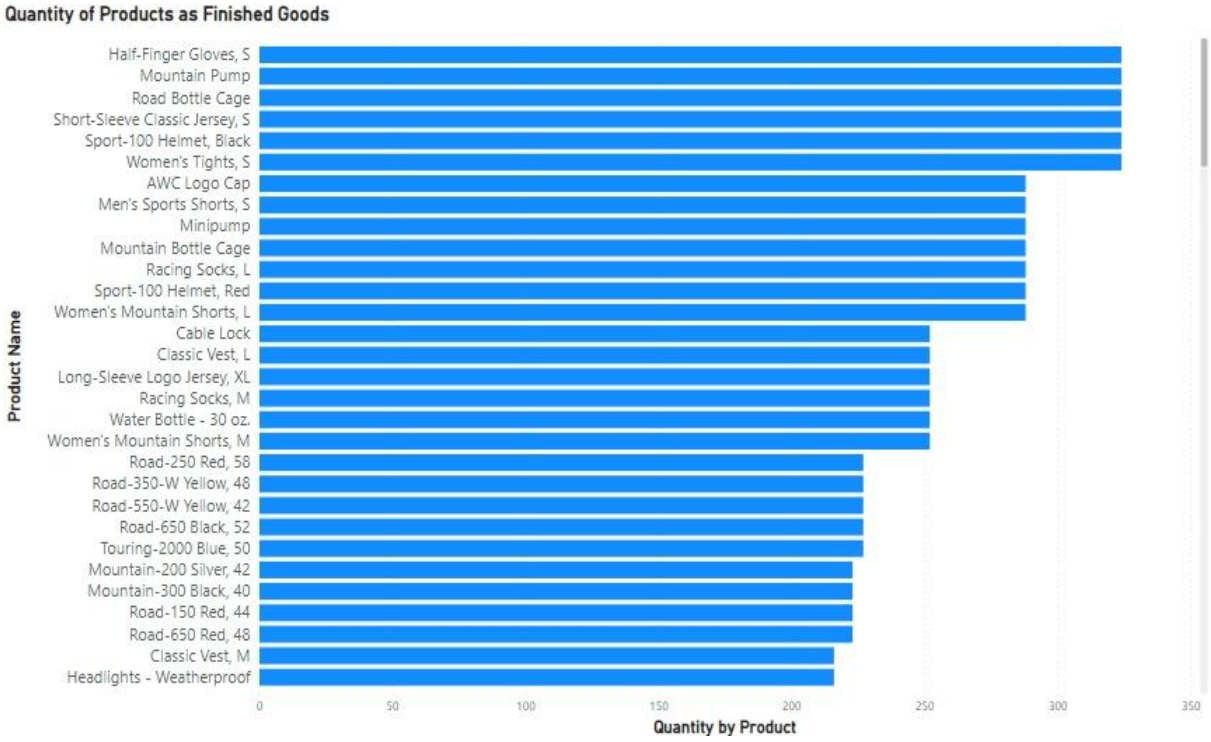
Here we see that quantity of bicycles and accessories produced each year has progressively increased until 2013 for a total of 27652 but decreased in 2014 to reach 17975. Then, 2013 was the most productive year.

```sql
--Calculating total quantity in inventory across all production units

SELECT Inv.ProductID, Inv.LocationID, Loc.Name, Inv.Quantity
,SUM(Quantity) OVER (PARTITION BY Inv.LocationID) AS Total_Qty_by_Unit
FROM AdventureWorks2014.Production.ProductInventory Inv
JOIN AdventureWorks2014.Production.Location Loc
ON Inv.LocationID = Loc.LocationID
WHERE Quantity > '0'
ORDER BY Total_Qty_by_Unit DESC
```



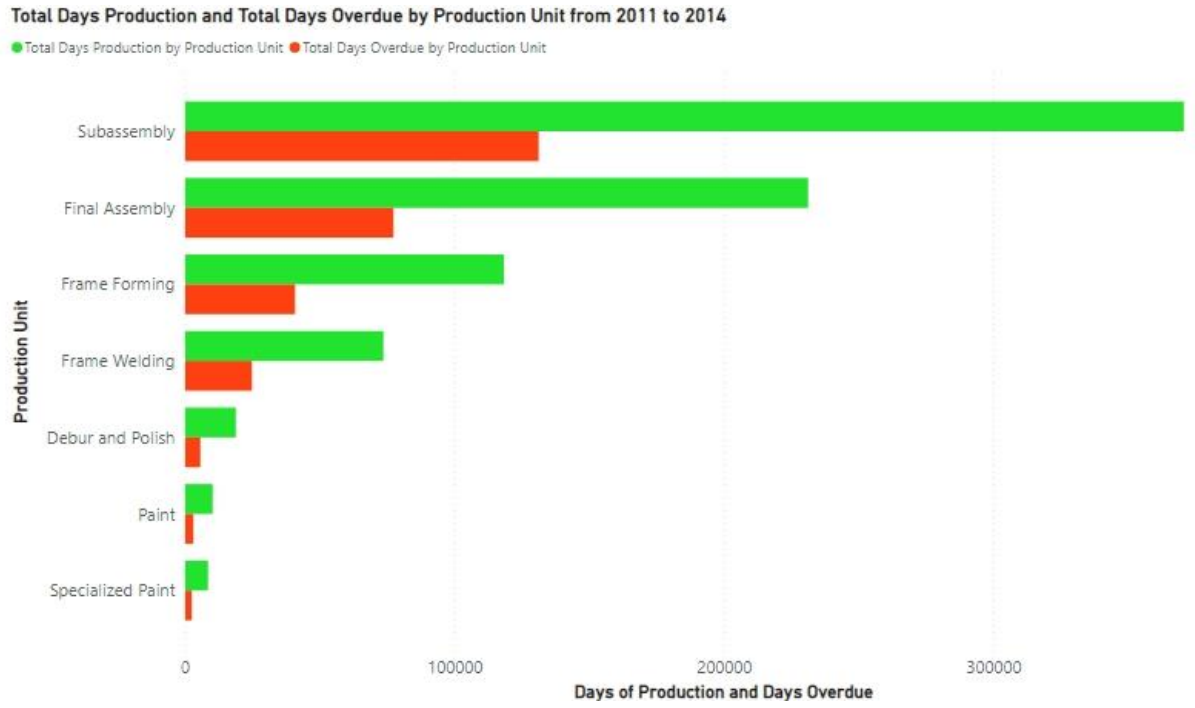
The production department consists of fourteen (14) sub-production units having specific roles. Subassembly has the highest production capacity with a total quantity of 95477 (29%) followed by Miscellaneous Storage that accounts for 83173 (25%). Tool crib has third highest production capacity with a total quantity of 72899 (22%).

```
--Showing products as finished goods ready for sale

WITH ProductInventory_CTE (ProductID, LocationID, Name, Quantity,
Total_Qty_by_Product)
AS(
SELECT Inv.ProductID,  Inv.LocationID, Loc.Name, Inv.Quantity
,SUM(Quantity) OVER (PARTITION BY Inv.ProductID) AS Total_Qty_by_Product
FROM AdventureWorks2014.Production.ProductInventory Inv
JOIN AdventureWorks2014.Production.Location Loc
ON Inv.LocationID = Loc.LocationID
)
SELECT ProductID
,LocationID
,Name AS Production_Unit
,Total_Qty_by_Product
FROM ProductInventory_CTE
WHERE LocationID = '7'
```



Finished goods storage has the sixth highest production quantity (17319). This unit production is the last step in the company's production chain and accounts for 152 products ready for sale. We notice that Half-Finger Glover, Mountain pump, Road Bottle Cage, Short-Sleeve Classic Jersey, Sport-100 Helmet and Women's Tights have 324 as highest quantity ready for sale.

```sql
--Showing Total Days of Production and Total Days Overdue by Production Unit from 2011
to 2014

SELECT Work.LocationID
,Name AS Production_Unit
,SUM(DATEDIFF(day, ActualStartDate, ActualEndDate)) AS
Total_Days_Production_by_Prod_Unit
,SUM(DATEDIFF(day, ScheduledStartDate, ActualStartDate)) AS
Total_Days_Overdue_by_Prod_Unit
FROM AdventureWorks2014.Production.WorkOrderRouting Work
JOIN AdventureWorks2014.Production.Location Loc
ON Work.LocationID = Loc.LocationID
GROUP BY Work.LocationID, Name
ORDER BY Total_Days_Production_by_Prod_Unit DESC
```



In this section, I calculated and compared total days of production and total days overdue by production unit from 2011 to 2014. To get total days of production I calculated the difference between ActualStartDate and ActualEndDate and found the sum for each production unit.

By contrast, some work orders were overdue. Then, to get those days overdue I calculated the difference between ScheduledStartDate and ActualStartDate and found the sum for each production unit.

This helps us understand that despite some days of production overdue, all of production units are generally within the time required to manufacture a product. In fact, if total days overdue were higher than total days of production for a production unit, this would mean that there is a problem within than production unit and would need to be fixed.

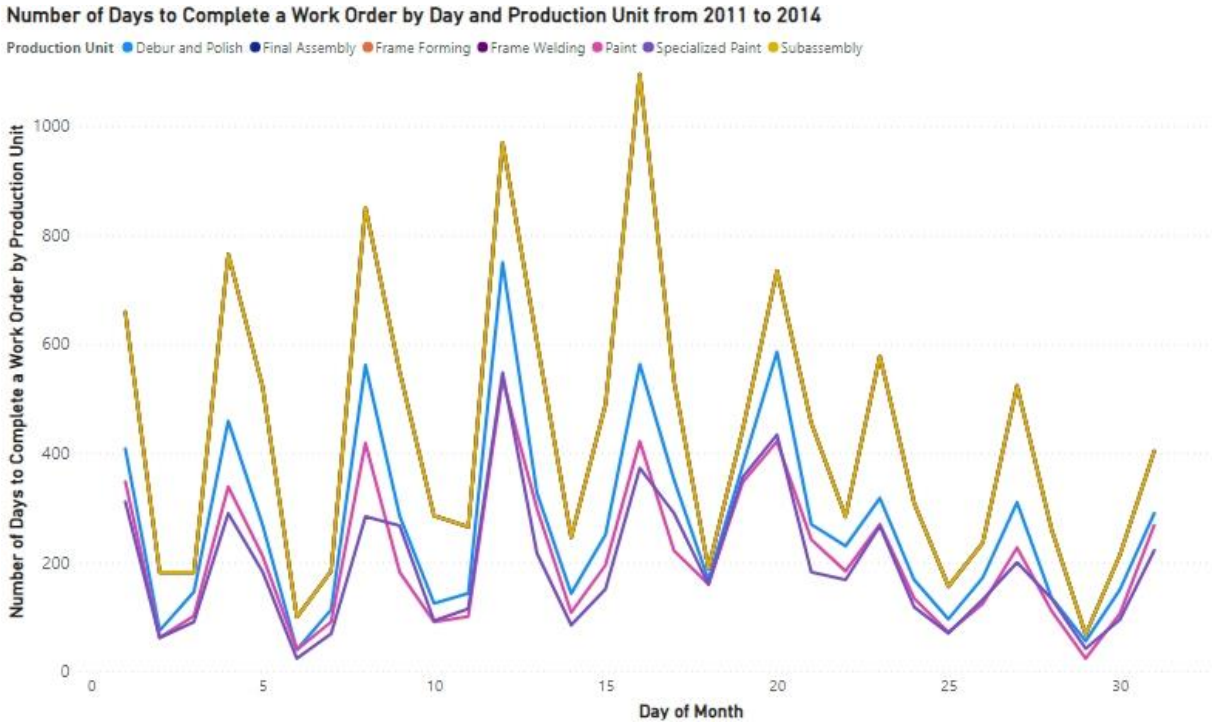
For example, remembering that subassembly has the highest quantity of production (95479), we notice on this graph that that production unit has also the highest days of production (370608) but also the highest days overdue (131271). We can see that other production units

display similar output. The highest the production quantity, the highest the total days of production and the total days overdue.

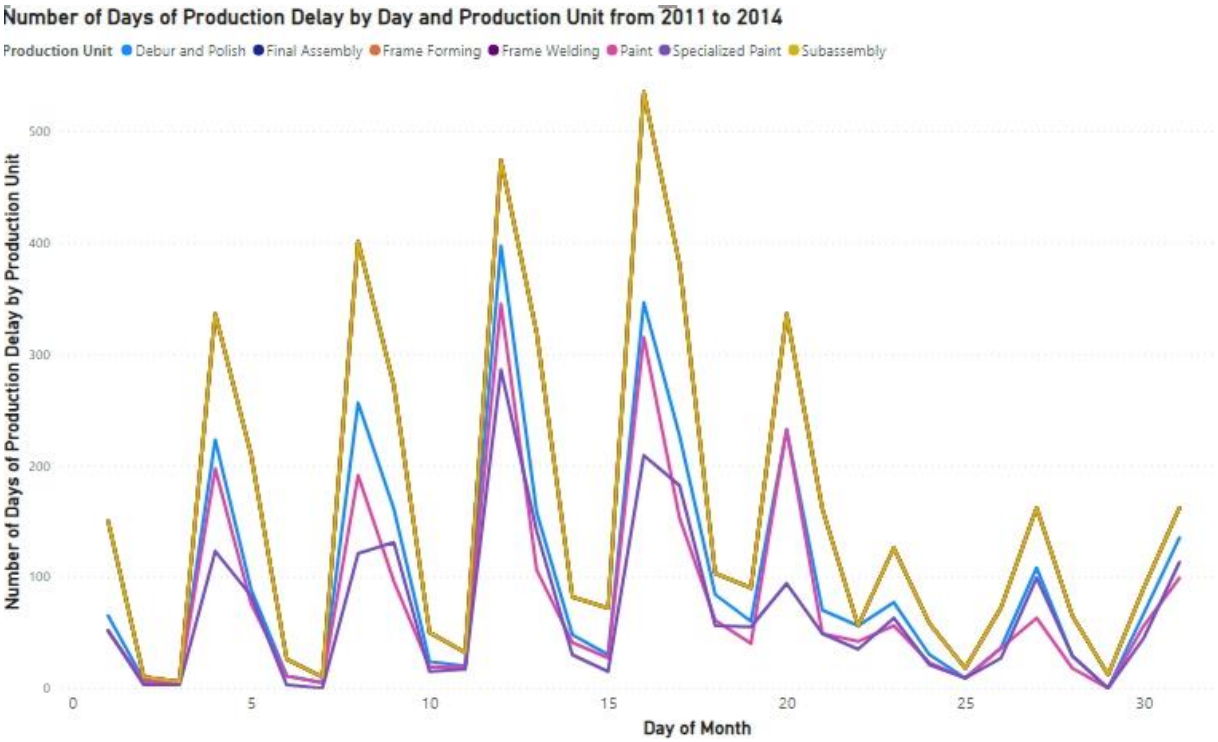
```
--Showing Number of Days of Production and Number of Days Overdue for Subassembly
Production Unit from 2011 to 2014
```

This section is a follow-up of the previous one, it is more detailed and helps us better understand the behaviour of each production unit when it comes to complete a work order. Results have been aggregated using days of month to better capture specific days with highest or lowest production activity.

```
SELECT Work.LocationID
,Loc.Name AS Production_Unit
,ScheduledStartDate
,ActualStartDate
,ActualEndDate
,DATEDIFF(day, ActualStartDate, ActualEndDate) AS
Number_of_Days_to_Finish_a_Work_Order
,DATEDIFF(day, ScheduledStartDate, ActualStartDate) AS Number_of_Days_Overdue
FROM AdventureWorks2014.Production.WorkOrderRouting Work
JOIN AdventureWorks2014.Production.Location Loc
ON Work.LocationID = Loc.LocationID
GROUP BY Work.LocationID, Name, ActualStartDate, ActualEndDate, ScheduledStartDate
```

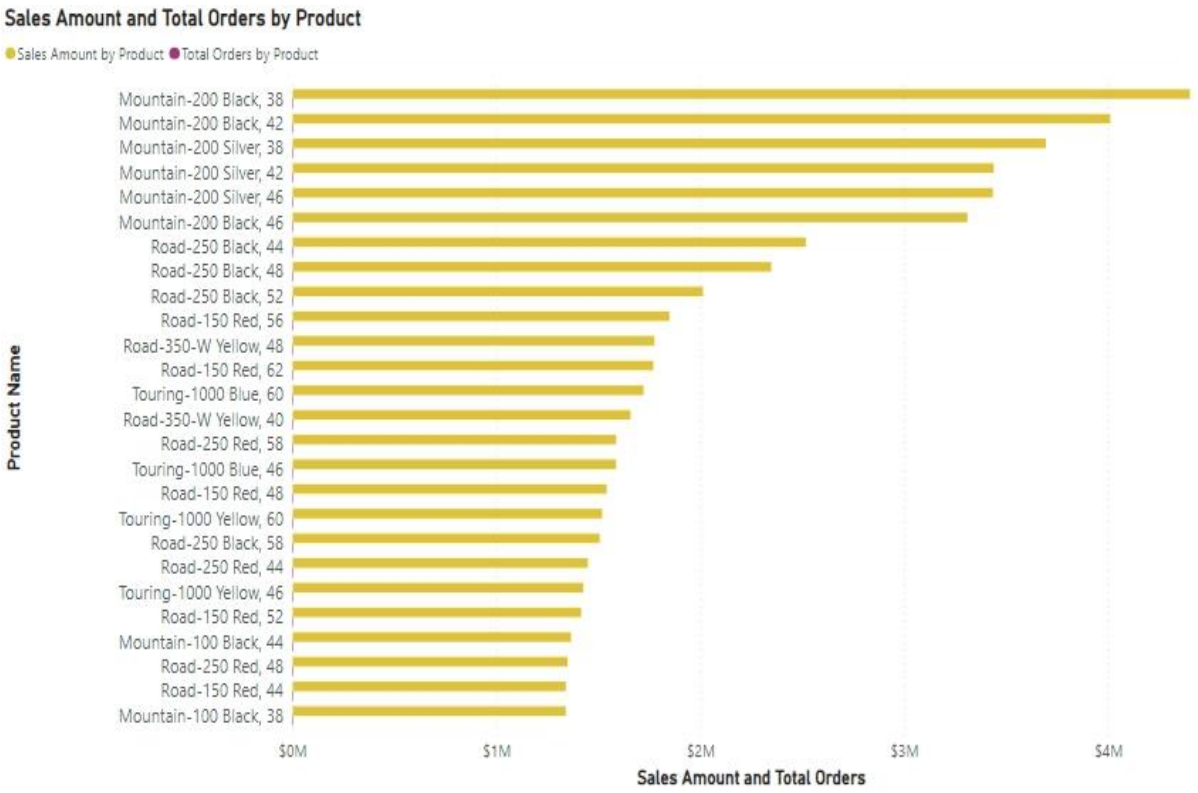


We notice that in this time series, subassembly is still top-ranked when it comes to the number of days to complete a work order by day. An important point to mention is that production activity is more intense in the middle of the month for each production unit



In this section again, subassembly has the highest number of days of production delay. This event occurs in the middle of the month also and for each production unit. This shows that there is a strong correlation between the days of production and the days of production delay.

```
--Calculating total orders and sales amount by product
WITH SalesOrder_CTE (LineTotal, ProductID, Name, OrderQty, Total_Sales_Amt_by_Product,
Total_Orders_by_Product)
AS(
SELECT ord.ProductID, ord.LineTotal, prod.Name, ord.OrderQty
,SUM(LineTotal) OVER (PARTITION BY ord.ProductID) AS Total_Sales_Amt_by_Product
,SUM(OrderQty) OVER (PARTITION BY ord.ProductID) AS Total_Orders_by_Product
FROM AdventureWorks2014.Sales.SalesOrderDetail ord
JOIN AdventureWorks2014.Production.Product prod
ON ord.ProductID = prod.ProductID
)
SELECT Name AS Product_Name
,FORMAT (Total_Sales_Amt_by_Product,'C','en-us') AS Total_Sales_Amt_by_Product
,Total_Orders_by_Product
FROM SalesOrder_CTE
GROUP BY Name, Total_Sales_Amt_by_Product, Total_Orders_by_Product
ORDER BY Total_Sales_Amt_by_Product DESC
```



In this section, we have a total of 274914 orders for 266 products and a total sales amount of $109.8 million. Mountain-200 Black, 38 has the highest sale amount that is $4.4 million and 2977 orders.

PowerBI Dashboard

To visualize an interactive dashboard containing these visuals, please see:

https://app.powerbi.com/reportEmbed?reportId=d27dc5ba-52aa-4782-b45f-b80a71d67510&autoAuth=true&ctid=c23476e1-b3f7-42ec-a7fc-971428e8b8ba

Analysis summary

- From 2011 to 2014, a total quantity of 335974 have been produced at Adventure Works Cycles
- 2013 is most productive year with a total quantity of 27652 products produced
- Subassembly has the highest production capacity (95477) among all of the production units
- Subassembly has the highest number of days of production (370608) and also the highest number of days of production delay (131271). These delays may have some negative impacts on production in general and it is the responsibility of the management to fix that.
- $109.8 million were generated through 274914 orders for a total of 266 products