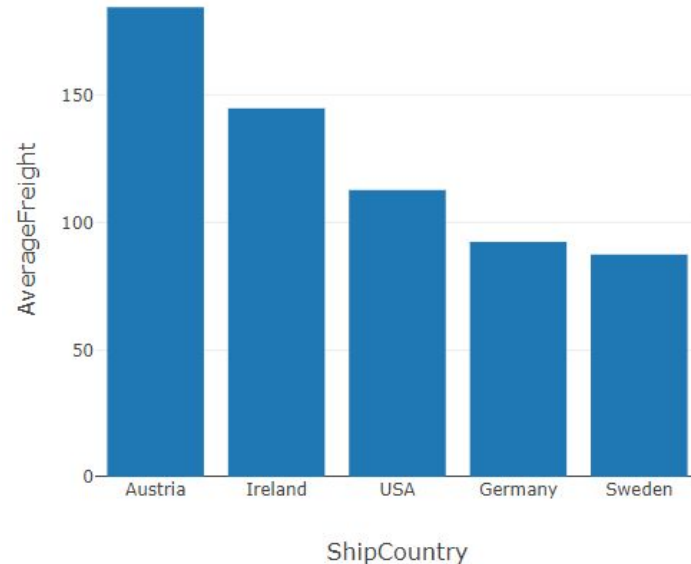
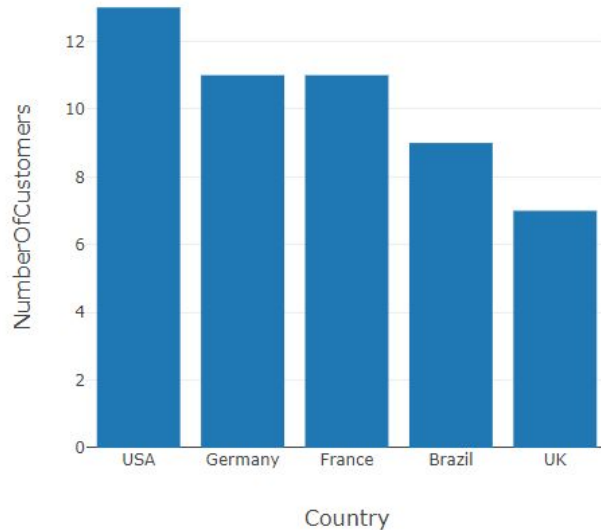


Northwind Sales and Product Analysis

Matt Calvert

Shipping Considerations

3 of the top 5 countries with the most customers are in Europe, and 4 of the top 5 countries with the heaviest average shipments are in Europe. It might be worth exploring a European distribution center to minimize shipping costs.



Biggest Spenders

The top 3 spenders have each spent twice as much (\$115K average) as the fourth biggest spender (\$57K). These customers should have a dedicated Account Representative to maintain a positive relationship with the customer.

```
1 %sql
2 -- Which customers are the biggest spenders in terms of total amount spent on orders?
3 -- List only companies with a total spending of greater than 20,000.
4 SELECT CompanyName, sum(UnitPrice*Quantity) AS TotalSpending
5 FROM orders
6 INNER JOIN customers
7 ON orders.CustomerID = customers.CustomerID
8 INNER JOIN order_details
9 ON orders.OrderID = order_details.OrderID
10 GROUP BY CompanyName, ContactName
11 HAVING TotalSpending > 20000
12 SORT BY TotalSpending DESC;
```

▶ (4) Spark Jobs


▶  _sqldf: pyspark.sql.dataframe.DataFrame = [CompanyName: string, TotalSpending: double]

Table ▾ +

	CompanyName	TotalSpending
1	QUICK-Stop	117483.390000000001
2	Save-a-lot Markets	115673.39
3	Ernst Handel	113236.68
4	Hungry Owl All-Night Grocers	57317.390000000001
5	Rattlesnake Canyon Grocery	52245.9
6	Hanari Carnes	34101.15
7	Folk och få H8	32555.55

Sales Pipeline

Two customers on our customer list have never placed an order. This data should be continuously monitored and shared with the sales team to encourage follow through on customer engagements. A monthly dashboard for the sales team might be helpful.

```
1 %sql
2 -- List customer IDs of customers who have never placed an order.
3 SELECT customers.CustomerID, orders.CustomerID
4 FROM customers
5 LEFT JOIN orders
6 ON customers.CustomerID = orders.CustomerID
7 WHERE orders.CustomerID IS NULL;
```

► (2) Spark Jobs

►  _sqldf: pyspark.sql.dataframe.DataFrame = [CustomerID: string, CustomerID: string]

Table ▼ +

	CustomerID ▲	CustomerID ▲	
1	FISSA	null	
2	PARIS	null	