**Introduction**

The research team investigated the company’s data warehouse which contains years of historical sales data. The company has not leveraged the data warehouse throughout the years, so this research team was tasked with deriving actionable business insights to prove its value. Five business objectives were investigated to properly understand the value of the data warehouse and determine its return on investment, if any, and secure future budgeting for the system. This report reviews each business objective in detail and explores the findings and actions that can be implemented by intentional utilization of the data warehouse.

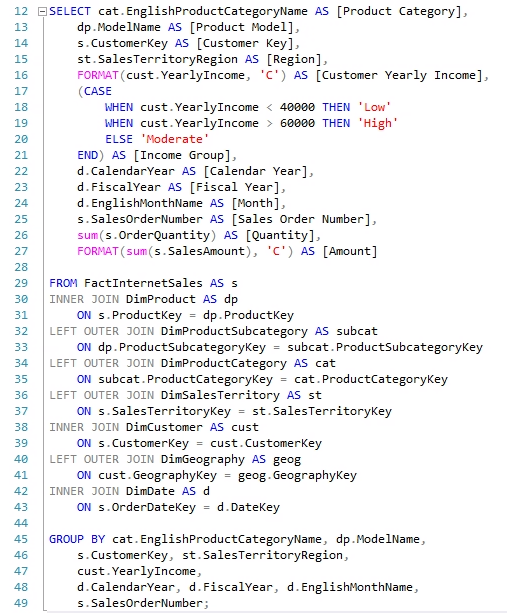
**Data and Assumptions**

The AdventureWorksDW2017 database includes 30 distinct tables that house pertinent data to our business. The tables are split between “Dim” tables, short for “Dimension”, which provide descriptive and contextual information for measurements found in the “Fact” tables. The area of focus based in this research centered around sales data broken up between internet sales and reseller sales, customer data including demographic information, product data that was split between categories and subcategories, promotional data, geographic data, and territory data. We were unable to modify any of the source data, so all data aggregations and manipulations were addressed directly in our queries. For example, the territory data was aggregated to depict the United Kingdom as a separate entity from Europe to reflect current circumstances. After gathering this data to a central location, Power BI was utilized to develop visualizations instead of any form of rollup. All queries were made using Structured Query Language (SQL) and executed in SQL Server Management Studio.

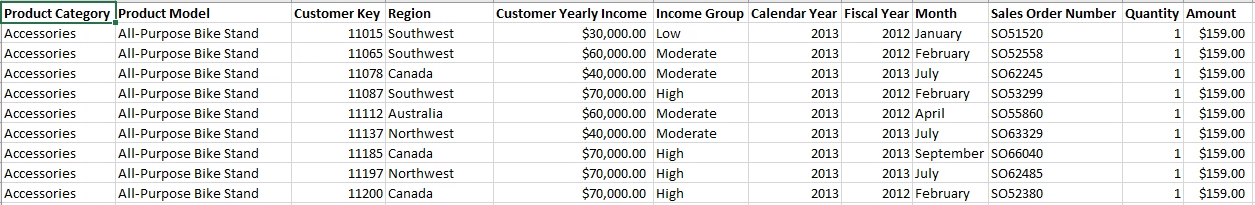
**Business Objective 1**

A direct approach was taken for the first objective. It was phrased as “Provide a detailed list of Internet sales with the following columns for the financial analyst team to review (Category, Model, CustomerKey, Region, IncomeGroup, CalendarYear, FiscalYear, Month, OrderNumber, Quantity, and Amount). Income group should categorize the people based on Low being less than 40,000, High being greater than 60,000, and the rest will be Moderate.” A base query was built by selecting the key columns. A CASE clause was used to create income categories for “High”, “Moderate”, and “Low”, and the appropriate filtering was done to display only the relevant information.

**SQL Query**

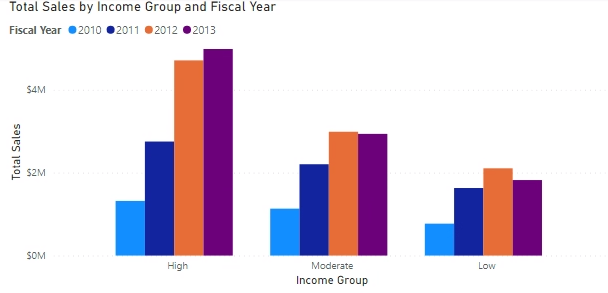


Below is a sample of the result set created by the query:



**Findings:**

As can be expected, internet sales were highest among our higher income customers. The highest sales growth can be seen in the ‘High’ income group between 2010 and 2013 resulting in $3.6MM of growth within the 4 year timespan. For all income groups, a net positive trend over time is seen, which indicates that sales have organically increased. The key takeaway from this round of insights is that we are seeing a positive trend in sales without the company’s awareness behind it. Knowing what we know now, we can be intentional with our marketing campaigns and set achievable sales targets to boost these margins even more. Further investigation on this took place in the fourth business objective regarding promotions, but additional data regarding advertising strategies would be necessary to conclude deeper insights.

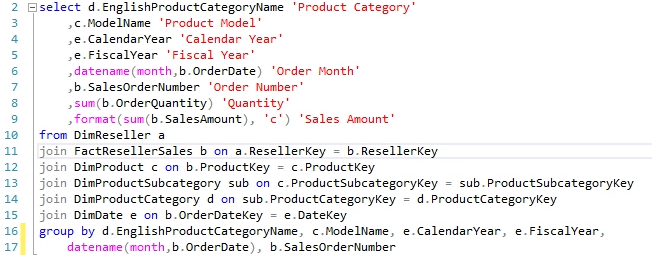


**Figure 1: Visuals depicting internet sales among income groups over time**

**Business Objective 2**

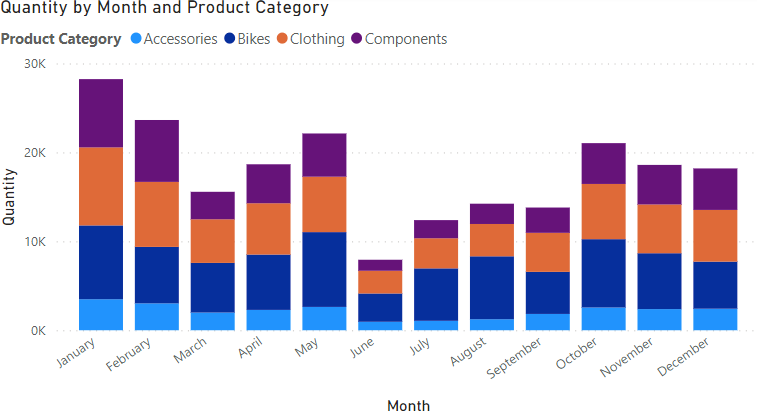
The second objective was similar to the first, phrased as “Provide a similar analysis for Reseller sales with the following columns (Category, Model, CalendarYear, FiscalYear, Month, OrderNumber, Quantity, Amount).” After creating a query with the base columns and adding proper formatting, a complex set of joins were used to align the proper data sets together. Grouping was used to create the calculations in the preferred portions.

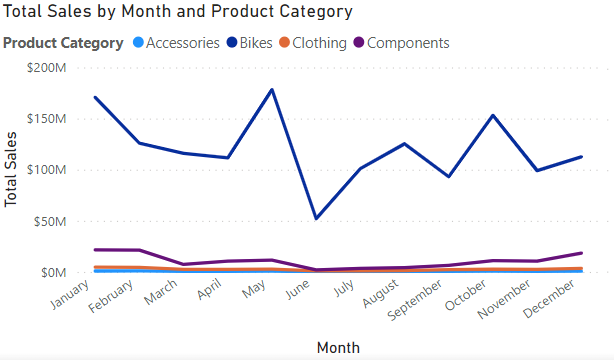
**SQL Query**

****

**Findings:**

Results from this query highlights our top selling products by resellers and also shows reseller sales trends over the calendar year. It was found that bikes are the company’s top selling product. Strongest sales are seen during Q1, with a sharp spike and decline in May and June, respectively. Sales trend upward for the remainder of the year, with a slight spike in October. It is recommended that the company should run holiday-centric promotions to capitalize on increased consumer spending during Q4 of each year, which is when most retailers expect larger sales volumes. Business objectives 1 and 2 can be leveraged with business objective 4 to drive further insights on how to implement the most effective strategy around marketing and promotions to help enhance sales.





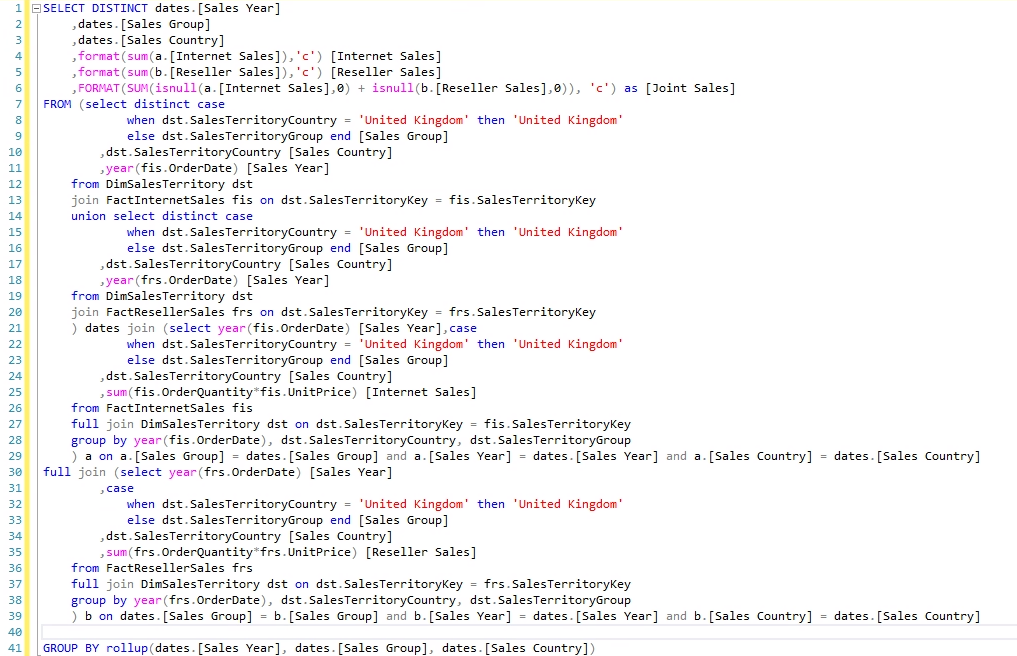
**Figure 2: Sales throughout the calendar year by product category**

**Business Objective 3**

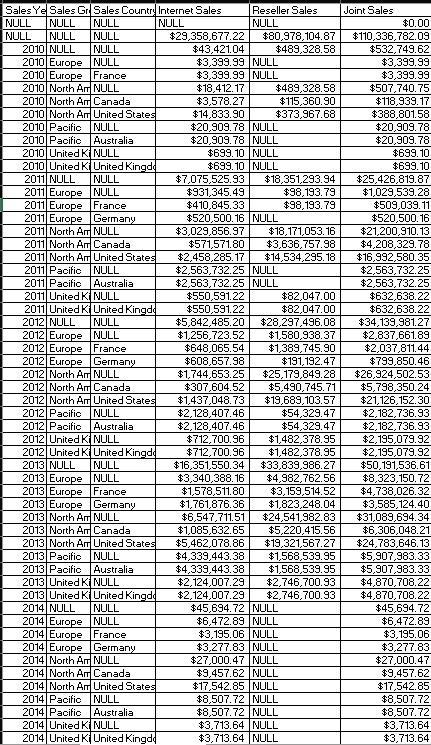
Business objective 3 was proposed to us as “Show the total sales (overall) by year rolled up by the Territory group and country. A special request is that the United Kingdom is no longer part of Europe and management wants to see their totals as a separate Territory group. You cannot modify the data, so you will need to address this request in your query.” The challenge in this query was determining how to properly separate out the year alone while also providing the country and the region. We began by writing the query to draw in data from both the Internet

sales and Reseller sales, with special focus on linking these two groups on the common grounds of Sales Group and Sales Country. Utilizing a CASE method to isolate the United Kingdom as its own country and region we were then able to create a complex union that made the rollup possible for Sales Year, Sales Group and Sales Country. This was then paired with a simple set of aggregation for Internet Sales, Reseller sales and Joint Sales. Previous attempts at other subquery joins had faltered with the rollup, so this approach was deemed best. For a quick assessment of aggregated sales, the team pulled data on an annual basis by sales group and country, subtotaling by internet and reseller customers for added context. The below query accomplished this goal.

**Query Code**

****

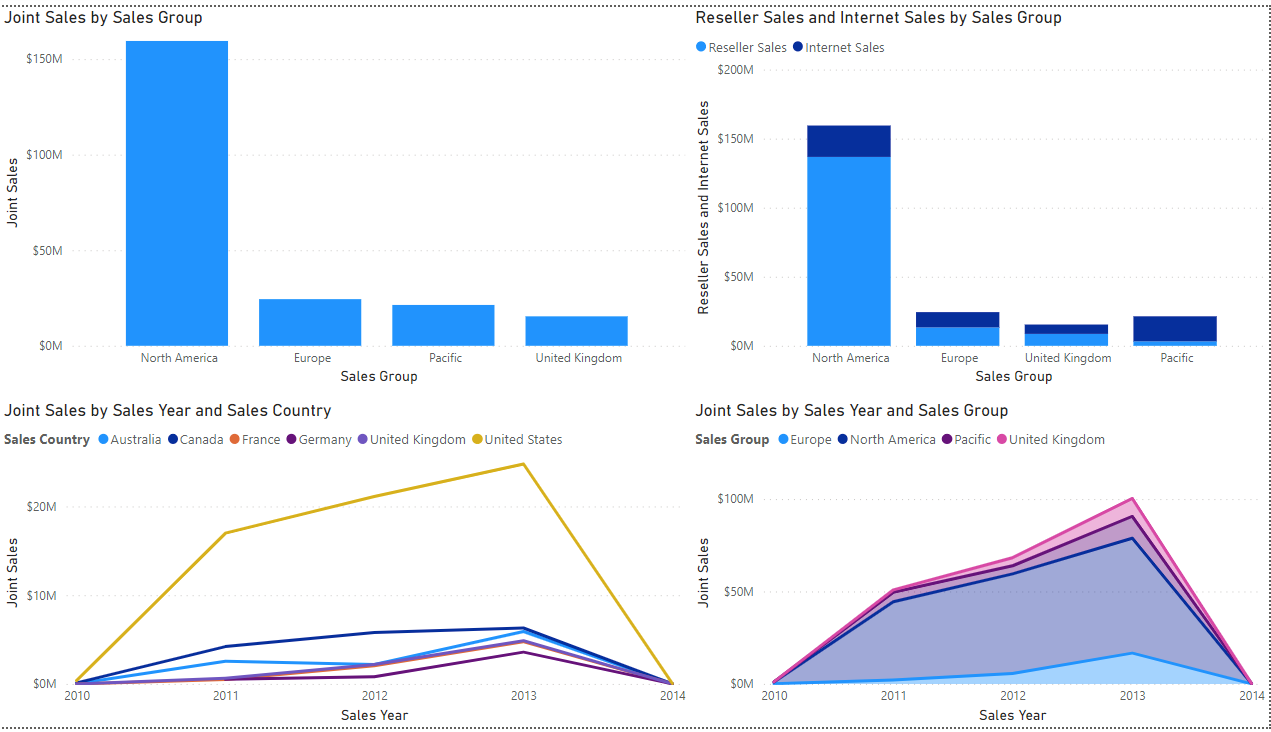
Below is a sample of the result set created by the query:



**Findings:**

As seen in the table, the rollup itself is useful for comparing the different countries' sales, and the different selling groups and their sale type. Some useful information derived is that North America is our largest customer, and that Reseller sales tend to out pace Internet sales. This was surprising, and indicates that we would benefit from more advertising campaigns for our Internet sales. Additionally, either due to the data or demand, our sales have increased over the last few years, with a jump of $24MM from 2011-2013. Going forward, special attention should be given to Europe, the UK, and the Pacific Sales groups, as they are significantly underperforming with only $62MM in joint sales over the last 3 years compared to North America’s $159MM.

Strong relationships with retailers lead to high reseller sales in North America. Internet sales are relatively consistent across all regions, although splitting the UK from Europe dilutes that market. While overseas shipping costs are high, there is untapped market potential that the company should consider investing in.

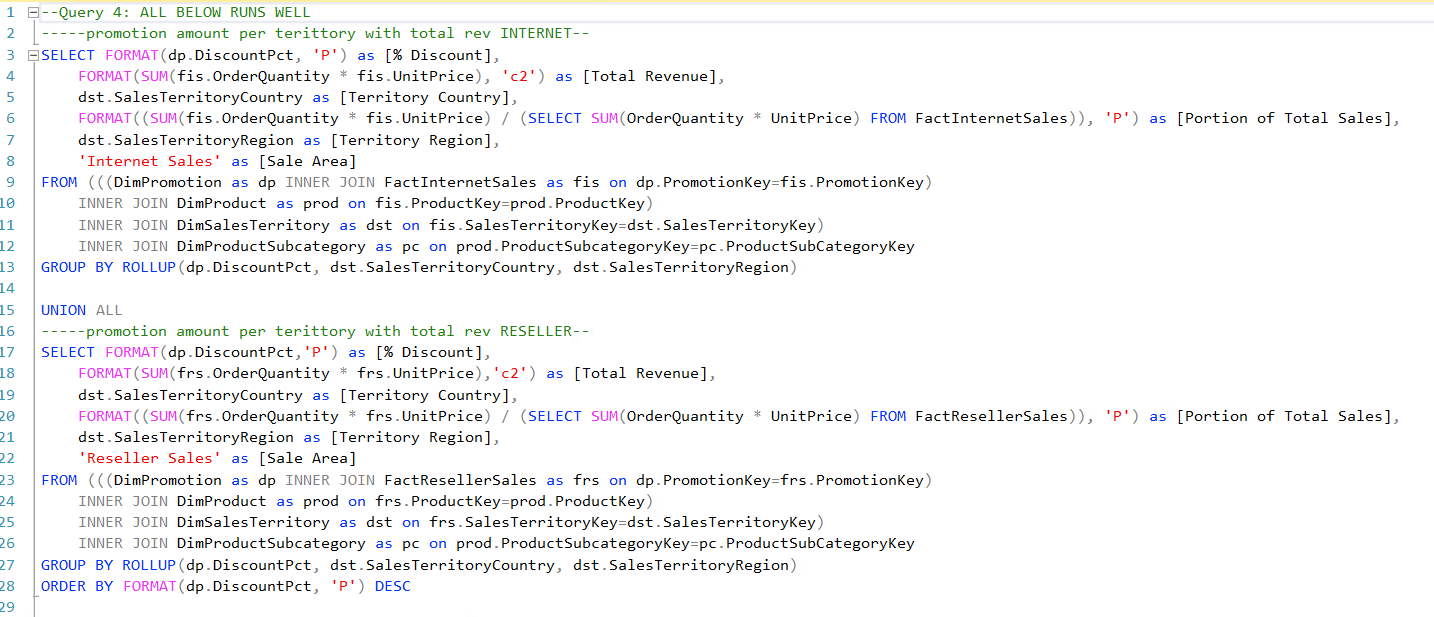
****

**Figure 3: Visuals depicting sales by region and by year.**

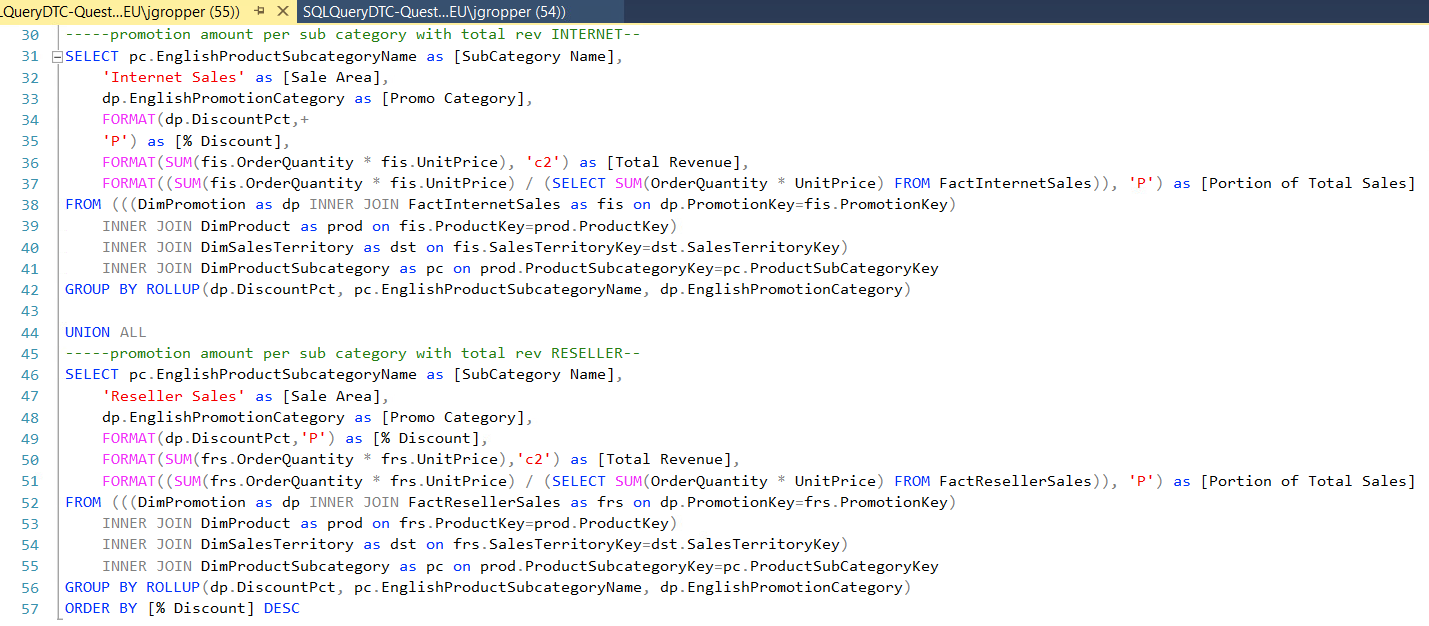
**Business Objective 4**

Next we were tasked with “an analysis of sales performance by Promotion. It would be interesting to see how different types of promotions drive sales (quantity and revenue), especially by product category or region”. When observing the promotion information of the fourth objective, we observed region and product category in two separate queries. The question that we were tasked with was to review performance of our promotions, especially by product category and region. We decided to go with a two query approach for this objective. We created separate queries for both Internet Sales and Reseller Sales for Promotion and region, then union joined them together. We did the same building and union join of the data for Reseller and Internet sales, but focused on promotion and product category. In order to determine full sales, we aggregated the quantity of each order multiplied by the amount. This enabled us to extract key summaries of information for the types of promotions and regions in one query, and the types of promotions with which category of product in another.

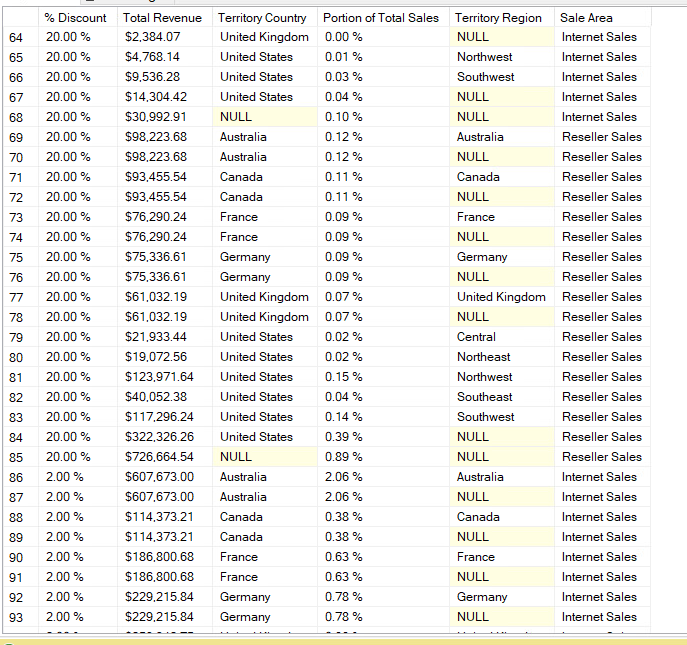
**Query Code 1**

****

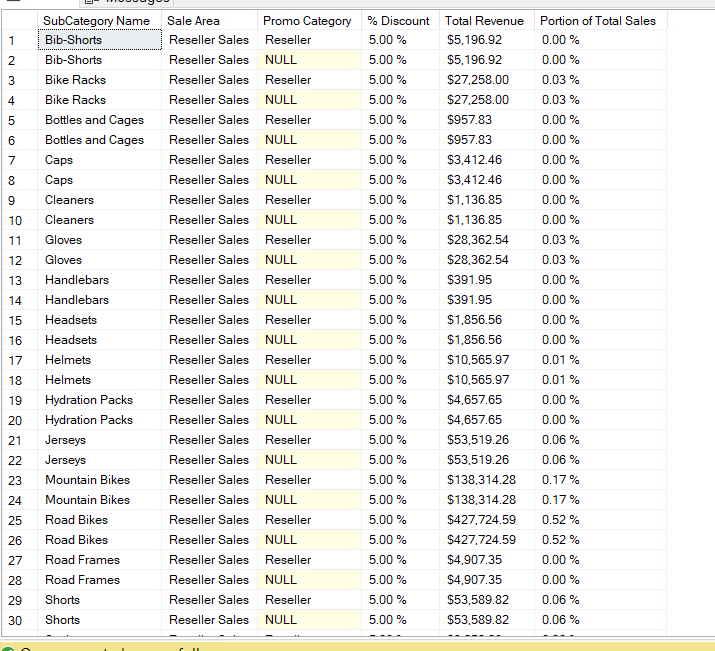
**Query Code 2**

****

Below is a sample of the result set created by the first query on Territories:

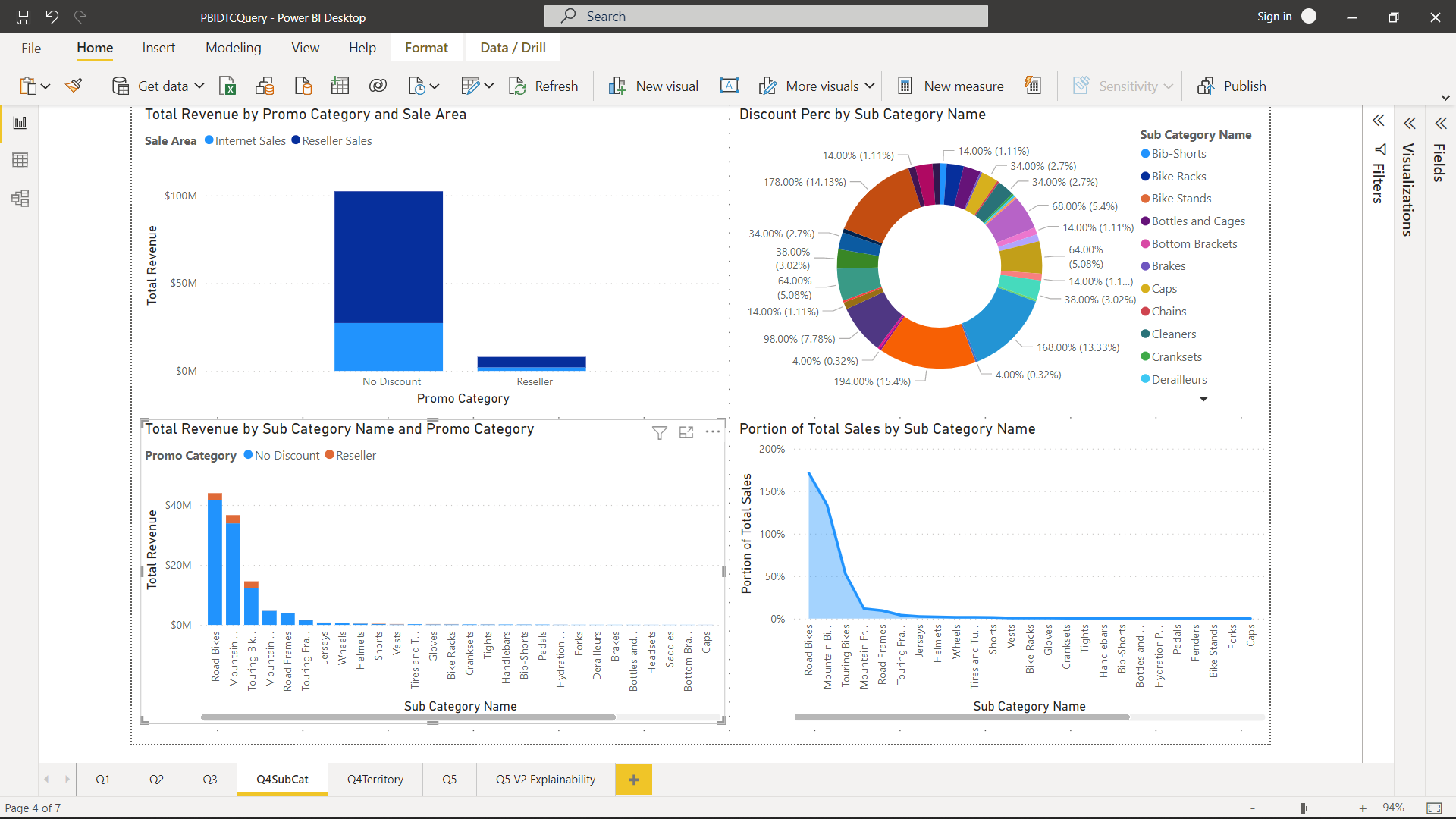


The second result set is a view of product categories on promotion:

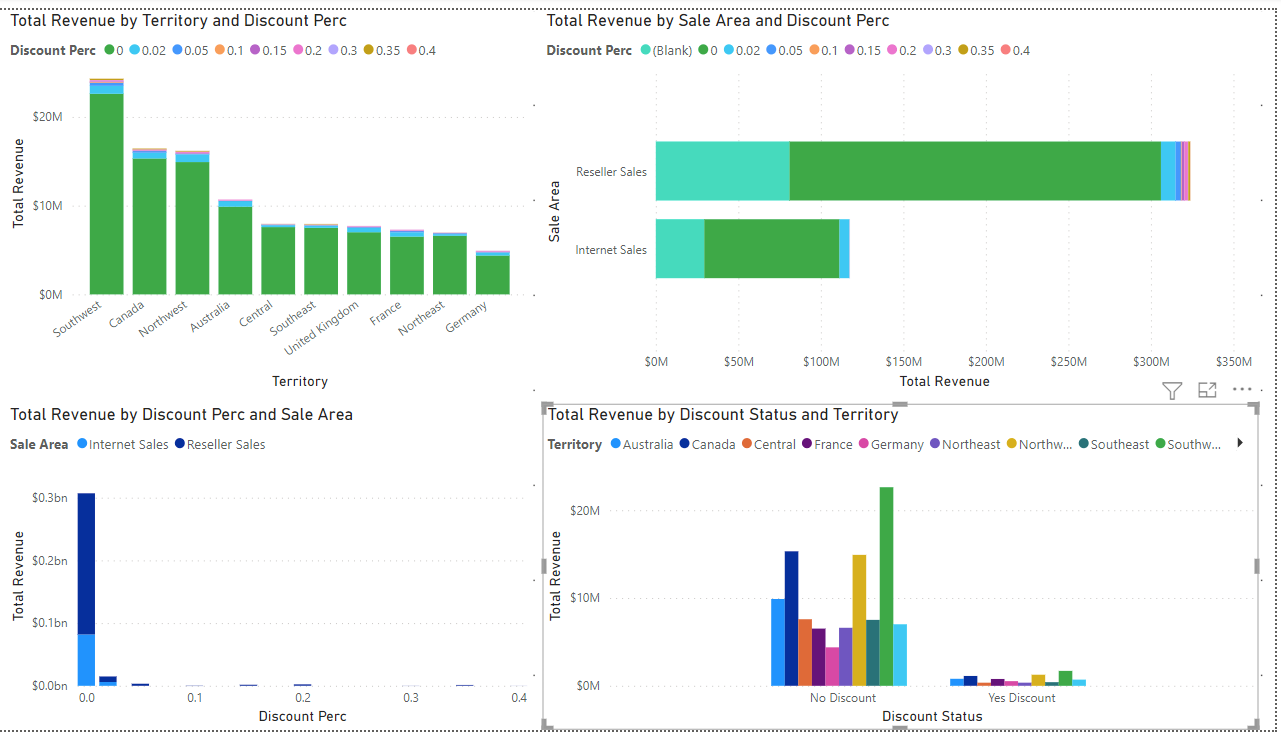


**Findings:**

For our distinct findings we were intrigued to discover that so little of our total sales are from promotional or discounted sales. Our sales show that $307MM is sold without a discount and $134MM is sold with a discount. This tells us that we have room for more sales potential if we marketed our discounts more effectively, especially considering that many of our discounts exceed 20%. Unsurprisingly, the majority of our sales are the physical bikes themselves, with few accessories being sold over time with discounts. We believe this could be improved upon as well, with better deals and discounts offered for products bought in tandem. When our attention is focused on the region and territory of sale, we found that Northwest and Southwest regions have the largest utilization of discounts with $1.7MM and $1.3MM, respectively. With our Southwest region having the most sales, and the most non-discounted purchases. Northeast, Southeast, and all of Europe do not take full advantage of our discounts with $25MM sold at full price and $2MM sold with a discount. This may be due to how unevenly our discounts are spread. It is shown that more discounts correlate with more sales, so we must bolster our campaigns in those three markets to increase sales.



**Figure 4: Visuals depicting revenue and sales by promotion**

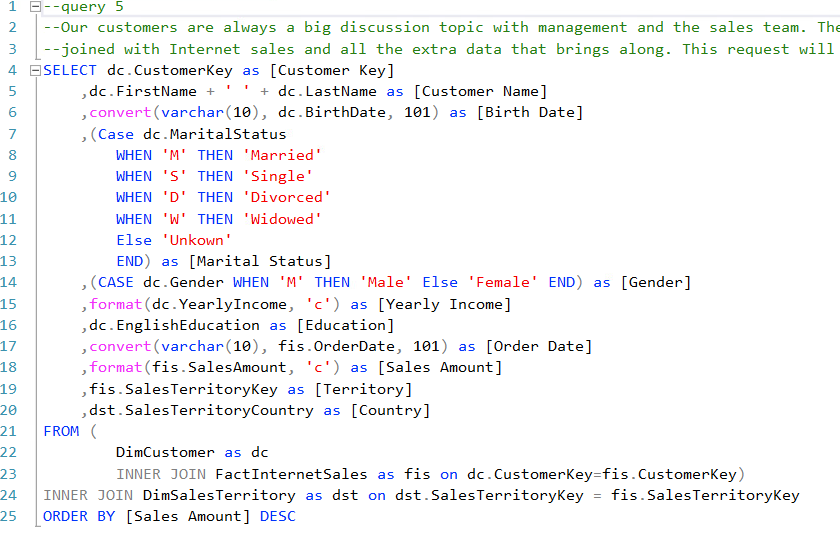


**Figure 5: Visuals depicting revenues by discount and region**

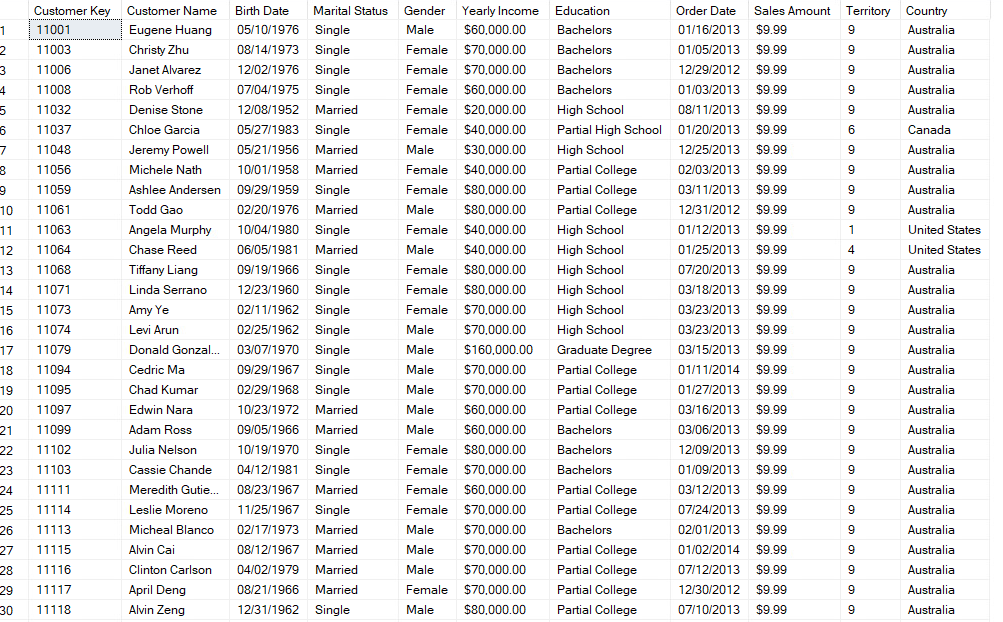
**Business Objective 5**

For the fifth and final question we were focusing specifically on our customer demographic data, joining it with the internet sales to discover trends, with an open ended request to find high quality analysis. We joined the customer data with the internet sales data and the territory sales data. We then proceeded to clean the data up with CASE When and some formatting. The proper joins allowed for the totality of the useful data to be gathered from fact and dimension tables.

**Query Code**

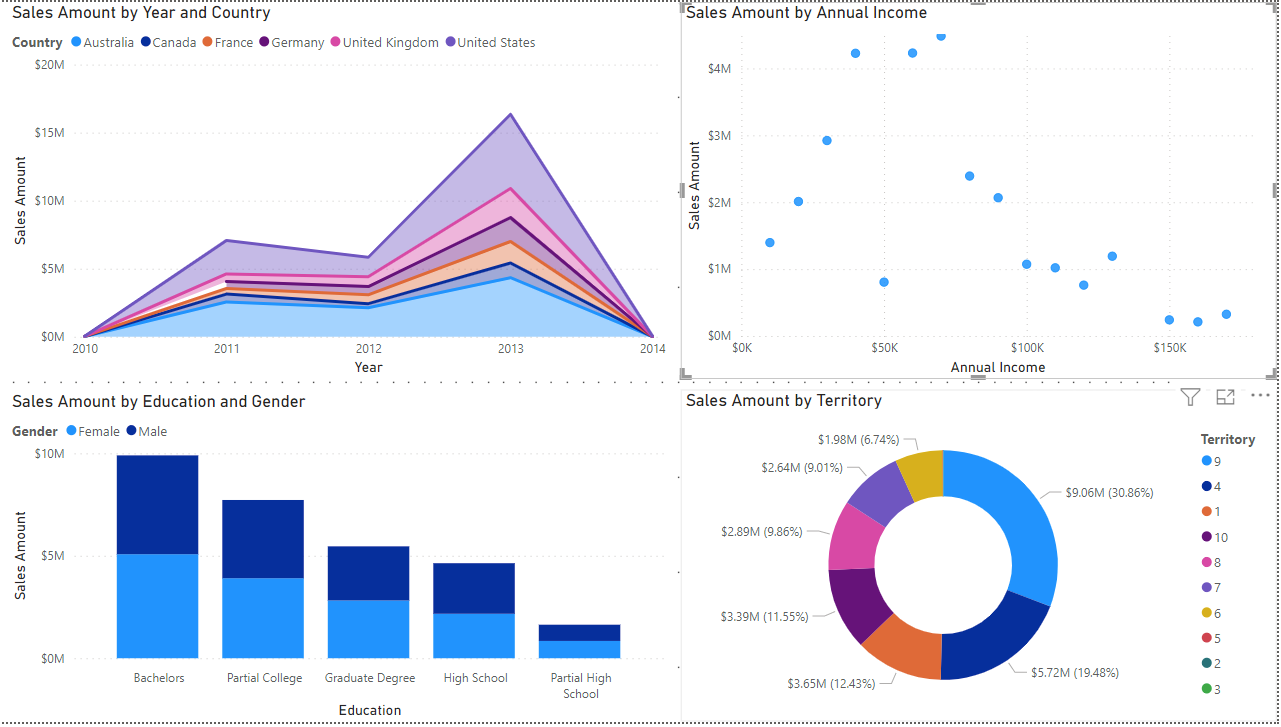
****

Below is a sample of the result set created by the query:

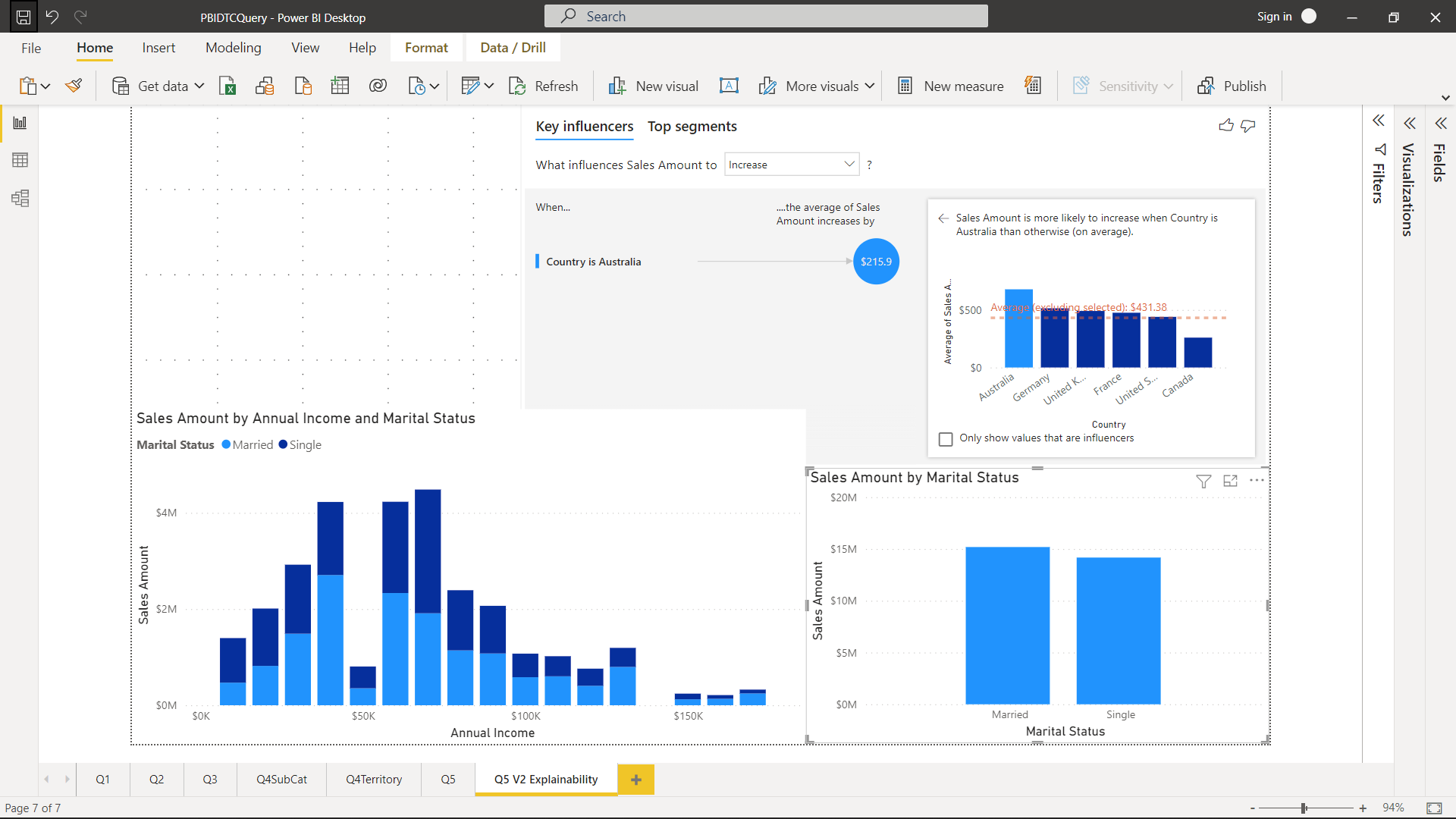


**Findings:**

We observed that the higher the educational attainment, the higher the sales. Customers with Bachelor's degrees purchased $1MM more than partial college customers. However, marriage status does not seem to impact sales, with about $14MM in sales for individuals identified as Single or Married. We also found that an increase in annual income from our customers does not always guarantee an increase in total sales. The majority of our purchases come from customers with annual income less than $100,000 with only about $4.5MM in sales sold to those outside of that income. This feels odd, and we think it would be well within the company’s interest to work to advertise more to upper income individuals to gather more business from that group with more purchasing power.



**Figure 6: Sales trends over time, country, education, gender, and annual income.**



**Figure 7: Sales amount by annual income and marital status**

**Conclusion and Future Opportunities**

The data warehouse has become a black box to the company, but the key findings derived from the business objectives discussed throughout this research should be evidence enough to warrant further investment of the system. With further maintenance of the system, the current data can be utilized even further, as seen with promotional strategies discussed in business objectives 1, 2, and 4. There are data types and relationships that should be updated so that the database can be analyzed more efficiently. Further, more data about our customers and their purchasing habits could be collected through the creation of new tables. Utilizing this could help us tap into millions of dollars in potential revenue, projecting our current growth as sustainable. Without proper data utilization, this company can’t expect to maintain its current annual growth rate in a highly competitive market, where all companies employ high quality analytics models.