**Assignment #5: Analyzing Data Cubes using Pivot Tables**

|  |
| --- |
| **Submission Instructions**  **Due:** **Friday, 3/23/2018 midnight**   * Please complete and submit the assignment as **a word or pdf** file through **Canvas>Assignments>To-Dos**.   + There is no answer sheet for this assignment.   + You can copy and paste the Pivot Tables directly from Excel. (If the pivot table is too big, it is okay to copy first 10 rows of the table)   + You do not need to submit the Excel file. * If you do not follow the instructions, your assignment will be counted late.   + Late Assignment policy: Same as before.   **Evaluation**  Your submission will be graded based on **(1) whether the answers provided for each question are correct** and **(2) the inclusion of a correct Pivot Table or Pivot Chart** (unless otherwise specified). |

You’ve been asked to help better understand sales patterns for Vandelay Interactive, an online retailer for clothing. They sell a variety of fine products, such as The Human Fund T-Shirt, the Schmoopie Bracelet, and Cougar 9000 Boxers.

The company has collected sales information for every line item (product sold, price, and quantity) for all 28,990 orders during 2017. Each order has multiple line items, for a total of 47,258 records. Vandelay Interactive also knows the referral source (how the customer got to the site), whether a Groupon was used, the product color and category, and the vendor that makes the product.

You will perform the quantitative analysis to answer each of the questions below. They can all be answered by constructing pivot tables and pivot charts in Excel. The Excel spreadsheet “**VandelaySales2017.xlsx**” has already been created for you. This is the single-table, joined data cube and is ready for analysis.

Here are some tips:

1. You will need to create pivot tables (or pivot charts) to answer all the questions.
2. For each answer, you must include the answer along with the Pivot Table/Chart – you can copy and paste it directly from the Excel spreadsheet.
3. You won’t need to create additional measures. The data in the table is all the data you’ll need. But you may apply formulas to the results of your Pivot Table analysis (for example, problems #4 and #5).
4. To answer some of these questions, you may need to perform several steps. For example, you may need to create one pivot table, and then use that information to filter data in a second pivot table.
5. The tab in the spreadsheet called “Legend” contains a brief description of each data label.

**Questions**

1. What are the five best-selling products based on revenue (i.e., total product price)?
2. What are the five worst-selling products based on quantity sold?
3. Which referral source brings in the highest revenue (i.e., highest total product price)?
4. What’s the average total quantity sold per category?  
   (Hint: Look at the total quantity sold for each category. Then use the Excel AVERAGE function to average those values by category in a separate cell ***outside*** the Pivot Table.)
5. What’s the average spending per customer (in dollars)? (Too many customers. So there is no need to include the Pivot table)  
   (Hint: Create a Pivot Table to get the total spending, i.e., total product price, for each customer\_id. Then use the Excel AVERAGE function to average those values by customer\_id in a separate cell ***outside*** the Pivot Table)
6. What category has the highest quantity sold in the summer months (June, July and August)?  
   (The answer should be for all three months together; don’t compute for each month separately.)
7. What is the best month for sales in 2017 based on total revenue (i.e. total product price)?  
   (You can include either a Pivot Chart or a Pivot Table that answers the question.)
8. The store is concerned that their Groupon promotion creates one-time customers that just use the Groupon and never come back. Is this true? Explain why you came to this conclusion.

(Provide evidence using a few customer examples from your Pivot Table and look at promo\_code – you don’t need to include the entire table.)

1. The most revenue is generated from products sold by which three vendors (by total product price)?
2. What are the three most popular products (by quantity sold) from those three vendors listed in #9?  
   (Your table should have three products in total aggregated across the three vendors, not three products from each vendor.)