Project 1.

For this project, you will work with students from Marketing. The **rosetta** database contains information about marketing campaigns to sell smart phones and tablets. You should have read-only access to it. Please read: <https://polylearn.calpoly.edu/AY_2015-2016/mod/forum/discuss.php?d=7015> for more information. The databases use a “star schema”. Fact tables contain information about different entities and usually have no foreign keys. Dimension tables, on the other hand, are comprised mostly of foreign keys and tell us how different entities relate. The database contains three dimension tables that are described next.

The **emailcmpaign** table contains a tuple for every time a customer does something with an e-mail they have received. The e-mails are “smart” and they can keep track of when the e-mail was opened, when a link inside it was clicked, when the customer unsubscribed from e-mails, and so on. **CustomerID** is a foreign key to the **customer** table (I have not created the f.k. because the data maybe a little dirty, i.e., it is possible to have references to customers that do not exist). The reference to the **campaign** table tells us which campaign this e-mail was from. There are two types of campaigns: loyalty and retention. In the **audience** table we describe the audience of the e-mail (e.g., single white female). In the **version** table we specify the version of the e-mail, where different people may get different versions of the same e-mail. The subjectLineKey field points to the subject line of the e-mail. The **DepolymentDate** is the date when the e-mail was sent. The event table describes the event: e.g., e-mail was opened. **EventDate** is the date of the event. **EventDateTime** stores date and time of event. **LinkName** is a key to the link table that contains internal information about the link that was clicked, while **URLKey** is a pointer to the table that contains the URL for the link.

The **device** table keeps information about the products that are sold, where only smart phones and tables are sold.

Lastly, the **product** table keeps information about product registration. It tells us which customer used what method to register what product on what day and when the product was actually purchased.

I have created a lot of indices, so your queries should run like the wind. Send me an e-mail if you need additional indices!!!

Submit information for 10 queries. For each query, include the English description and the relational algebra expression. The marketing students need to approve your 10 queries and they need to be complicated (e.g., some of them need to involve joins, grouping, and aggregates). You will need to create intermediate tables for most queries.

Submission is through PolyLearn, where only one team member needs to submit the project.

Ryan Sasur

GSB516

Assignment 3 – Requested SQL Queries

1. What is our registration rate of our products?
   1. (select count(\*) from product where purchasedatekey != 0 ) / (select count(\*) from product);
2. What is the income level of our customers, in relation to our different products?
   1. AVG income of customer for each product
3. Where are the most populated areas with customers?
   1. Max area of customers
4. Where are the most populated areas, in relation to customer tier?
5. What are the life cycles of our products? (Check the sales of our individual products over time)
6. What is the response rate to our email campaigns, in relation to segment?
7. What is the gender segmentation of our products?
8. What are our most popular products?
9. What is the early adoption rate of our devices?
10. When is the best time to send out our email campaigns?