Ancestry Product Analytics Homework Assignment:

Ancestry helps customers explore their family history through family tree creation and results of its DNA product. Since the launch of DNA product in 2012, there is a percentage of DNA customers who successfully became Family History Subscribers each year. From an optimization standpoint, there is not a clear ceiling on the cross-sell conversion rate and there may be upside potential. We believe there is an opportunity to turn every DNA customer into a Family History customer with the right experiences and value propositions over time.

Product Analytics team is interested in supporting Product Management in forming the cross-sell strategy analytically from DNA to ACOM subscription. To help answer the question, there is a sample dataset of users who purchased DNA product and their cross-sell activity.

We consider a registrant to be a Xseller (from DNA to Acom) when the following conditions are met Xsell_gsa == 1 and Xsell_day_exact <= 120.

We would like you to use this data set to help understand which factors best predict for Xseller, and offer suggestions to operationalize those insights to help Ancestry.

See below for a detailed description of the dataset (Link to dataset). Please include any supporting analysis or code written.

1. Define the problem: What is the cross-sell status today based on the dataset? For example, what fraction of the observed users cross-sell to subscription? Perform any cleaning, exploratory analysis, and/or visualizations to use the provided data for this analysis (a few sentences/plots describing your approach will suffice).

2. Discuss your analytics plan/steps to derive insights, identify opportunities, and recommend any optimization. How Ancestry might leverage the insights gained from the analytics deliverable to improve?

3. If you need to design an analytical framework to learn the impact of the above-mentioned insights iteratively, what experimentation method would you use to verify the recommendations:

- How would you set up the experiment? Test vs. Control or other.
- What data or information is necessary to determine the sample size and duration of the experiment? Which metrics would you plan to track that define performance of the test?
- What method would you use to analyze the test results? Why?
- Based on your method, when would you be able to say that we can go ahead launch the new product?

Column	Description
prospectid	Unique identifier for customers
ordernumber	Id for DNA order
ordercreatedate	Date DNA order is created; example '2016-02-08'
regtenure	Registrant tenure of the customer
customer_type_group	Customer type
dnatestactivationdayid	Date DNA test is activated. Customers need to activate their kit after ordering for testing
daystogetresult_grp	Number of days taken for result to be ready
dna_visittrafficsubtype	Dna Traffic visit channel, example: 'Paid Search Non Brand'
xsell_gsa	Gross Subscriber Addition
xsell_day_exact	Days between cross sell and subscription

Upon completion, please email to Sophine Zhang at szhang@ancestry.com. Please do not hesitate to reach out if you have any questions whatsoever.

Thank you!
Dorris O'Brien
Ancestry
Dorris.obrien@ancestry.com
801.705.7916