Collaborative Research Opportunity
Customer Lifetime Value for Business-to-Business Customers of a Leading B2B Supplier of Industrial Products and Services

January 18th, 2019
Introductions

Mary Purk
Executive Director
WCAI

Brandon Krakowsky
Research & Education Director
WCAI

Kayleigh Smoot
Associate Research & Education Director
WCAI

Company Representative A
Data Science Manager
Leading B2B Supplier of Industrial Products and Services

Company Representative B
Data Science Group Manager
Leading B2B Supplier of Industrial Products and Services

Company Representative C
Data Scientist
Leading B2B Supplier of Industrial Products and Services
The Wharton Customer Analytics Initiative (WCAI) is the world’s preeminent academic research center focusing on the development and application of customer analytics methods.

Through our broad network of top companies spanning a wide spectrum of sectors, WCAI has fostered opportunities for experiential learning and innovation in analytics for both students and faculty.
Questions?

To ask a question, click the chat button in the upper-right and type your question

• All questions will be answered live or via e-mail
Agenda

About the Company

Research Opportunities

About the Data

Research Proposal & Application Process

Q&A

Transforming Business Through Customer Data
A Few Operational Notes

This webinar is intended for researchers & scholars. It should not be published or presented without permission from WCAI.

Questions about our research programs or for press inquiries:

wcai-research@wharton.upenn.edu

To work with WCAI and the data sponsor, research teams must submit a proposal for approval. More details throughout this presentation.
About the Company
About the Company

- We are a Leading B2B Supplier of Industrial Products & Services
- What makes B2B special?

<table>
<thead>
<tr>
<th><strong>Unique Channels</strong></th>
<th><strong>Wide Customer Heterogeneity</strong></th>
</tr>
</thead>
</table>
| Products & services are provided to *large* and *small* B2B customers across a *variety of channels*, including:  
  - E-commerce  
  - E-mail  
  - Brick & mortar stores  
  - Telesales  
  - Field sales reps | B2B customers tend to make *recurring orders* with a wide heterogeneity in customers |
About Us

• Who We Are
  - **Company Representative A**: Data Science Manager
  - **Company Representative B**: Data Science Group Manager
  - **Company Representative C**: Data Scientist

• Our Team
  - We manage a team of 10 data scientists with a background in computer science, statistics and operations research
  - Together, we work on projects on *marketing, economics* and *healthcare*
About Us – Tools & Technology

- **Theory**: AI, deep learning, machine learning and game theory
- **Languages**: R, Python, and Java
- **Development**: R Studio, Jupyter Notebook, PyCharm
- **Deployment**: Docker, Microsoft Azure and Amazon AWS
Research Opportunities
Research Opportunities

We are seeking proposals in the area of customer lifetime value (CLV) research for B2B customers in a non-contractual setting with continuous opportunities of transactions

• Measuring salesforce effectiveness
• Predicting the effect of pricing on CLV
• Exploiting unique channel behavior in the prediction of CLV
• Incorporating other customer characteristics (e.g. firm characteristics, attitudes) and marketing activities into CLV predictions
• Accounting for nonstationarity in customer behavior (e.g. latent customer characteristics that can change over time)
Research Opportunities

We are seeking proposals in the area of customer lifetime value (CLV) research for B2B customers in a non-contractual setting with continuous opportunities of transactions

• Relaxing traditional assumptions about the distribution of purchases and drop-out rates
• Applying new machine learning methods to CLV prediction including multitask learning and deep learning
• Developing Bayesian non-parametric estimation methods
• Understanding the trade-off between model accuracy and computational overhead
About the Data
About the Data

- The dataset includes sales data for 10,000 non-contractual US customers, along with website clickstream data and e-mail/marketing data.
• The dataset includes sales data for 10,000 non-contractual US customers, along with website clickstream data and e-mail/marketing data.

• B2B customers can be vastly different
  - For example, we may observe small business owners who order products online, every few months
  - We may also observe owners of multiple franchise stores who order products in bulk over the phone, multiple times a month
About the Data

• The dataset includes sales data for 10,000 non-contractual US customers, along with website clickstream data and e-mail/marketing data.
• B2B customers can be vastly different
  - For example, we may observe small business owners who order products online, every few months
  - We may also observe owners of multiple franchise stores who order products in bulk over the phone, multiple times a month
• The data has been collected over a two year observation window
A Single Customer in the Data - Customer Scenario A

• We observe *registered* Customer A visiting the company’s website
A Single Customer in the Data - Customer Scenario A

- We observe *registered* Customer A visiting the company’s website
- Then we observe email campaign data sent to *registered* Customer A

Customer A

Website Visit, 08/19/2018

Marketing Email Sent & Opened, 08/19/2018

Today
A Single Customer in the Data - Customer Scenario A

- We observe *registered* Customer A visiting the company’s website
- Then we observe email campaign data sent to *registered* Customer A
- We observe sales transactions as *registered* Customer A makes purchases

Customer A

Website Visit, 08/19/2018
Marketing Email Sent & Opened, 08/19/2018
Items purchased, 08/19/2018

Today
A Single Customer in the Data - Customer Scenario A

• We observe *registered* Customer A visiting the company’s website
• Then we observe email campaign data sent to *registered* Customer A
• We observe sales transactions as *registered* Customer A makes purchases

And we observe all of this for **10,000 customers**, over a **two year period**
A Single Customer in the Data - Customer Scenario B

- Here, we first observe registered Customer B visiting the company’s website
A Single Customer in the Data - Customer Scenario B

- Here, we first observe *registered* Customer B visiting the company’s website
- Then we observe them making a purchase by calling to place their order
A Single Customer in the Data - Customer Scenario C

- In this scenario, we observe *registered* Customer C receive a marketing email.
A Single Customer in the Data - Customer Scenario C

- In this scenario, we observe *registered* Customer C receive a marketing email
- And then visiting a store to place their order

Customer C

Marketing Email Sent & Opened, 11/22/2018
Visit store to make purchase, 11/24/2018

Today
Data Structure

Registered Customers

- Sales Transaction Data
- Web Traffic Data
- Email Marketing Data
Data Structure

Registered Customers

Sales Transaction Data

Web Traffic Data

Email Marketing Data
Sales Transaction Data

- Each row in this table represents the *purchase of a particular product* (Product SKU)

<table>
<thead>
<tr>
<th>ID</th>
<th>Sales Order ID</th>
<th>Purchase Date</th>
<th>Customer ID</th>
<th>Product SKU</th>
<th>Product Category</th>
<th>Sales Revenue</th>
<th>Sales Cost</th>
<th>Profit</th>
<th>Currency</th>
<th>Purchase Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>111233</td>
<td>08/19/2018</td>
<td>1248657</td>
<td>65</td>
<td>HMH-DLU-JUM-DRE</td>
<td>550.00</td>
<td>450.00</td>
<td>100.00</td>
<td>USD</td>
<td>Call center</td>
</tr>
<tr>
<td>2</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>2</td>
<td>HMH-DLU-UIU-FRE</td>
<td>250.00</td>
<td>200.00</td>
<td>50.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>3</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>29</td>
<td>HMH-POU-AWE-SDE</td>
<td>15000.00</td>
<td>13000.00</td>
<td>2000.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>4</td>
<td>111235</td>
<td>11/15/2018</td>
<td>1234568</td>
<td>199</td>
<td>HMH-NHU-DFD-XDS</td>
<td>2500.00</td>
<td>2000.00</td>
<td>500.00</td>
<td>USD</td>
<td>Store</td>
</tr>
<tr>
<td>5</td>
<td>111236</td>
<td>11/17/2018</td>
<td>1234569</td>
<td>43</td>
<td>QEW-GFT-SAD-DEW</td>
<td>947.00</td>
<td>850.00</td>
<td>97.00</td>
<td>USD</td>
<td>Store</td>
</tr>
</tbody>
</table>

* Only a sample of data is represented. Awarded teams will have access to the full dataset.*
Sales Transaction Data

- Each row in this table represents the *purchase of a particular product* (Product SKU).

<table>
<thead>
<tr>
<th>ID</th>
<th>Sales Order ID</th>
<th>Purchase Date</th>
<th>Customer ID</th>
<th>Product SKU</th>
<th>Product Category</th>
<th>Sales Revenue</th>
<th>Sales Cost</th>
<th>Profit</th>
<th>Currency</th>
<th>Purchase Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>111233</td>
<td>08/19/2018</td>
<td>1248657</td>
<td>65</td>
<td>HMH-DLU-JUM-DRE</td>
<td>550.00</td>
<td>450.00</td>
<td>100.00</td>
<td>USD</td>
<td>Call center</td>
</tr>
<tr>
<td>2</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>2</td>
<td>HMH-DLU-UIU-FRE</td>
<td>250.00</td>
<td>200.00</td>
<td>50.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>3</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>29</td>
<td>HMH-POU-AWE-SDE</td>
<td>15000.00</td>
<td>13000.00</td>
<td>2000.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>4</td>
<td>111235</td>
<td>11/15/2018</td>
<td>1234568</td>
<td>199</td>
<td>HMH-NHU-DFD-XDS</td>
<td>2500.00</td>
<td>2000.00</td>
<td>500.00</td>
<td>USD</td>
<td>Store</td>
</tr>
<tr>
<td>5</td>
<td>111236</td>
<td>11/17/2018</td>
<td>1234569</td>
<td>43</td>
<td>QEW-GFT-SAD-DEW</td>
<td>947.00</td>
<td>850.00</td>
<td>97.00</td>
<td>USD</td>
<td>Store</td>
</tr>
</tbody>
</table>

- You can imagine that each purchase represents *an individual line item associated with a single order* (Sales Order ID).

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Sales Transaction Data

- Each row in this table represents the purchase of a particular product (Product SKU)

<table>
<thead>
<tr>
<th>ID</th>
<th>Sales Order ID</th>
<th>Purchase Date</th>
<th>Customer ID</th>
<th>Product SKU</th>
<th>Product Category</th>
<th>Sales Revenue</th>
<th>Sales Cost</th>
<th>Profit</th>
<th>Currency</th>
<th>Purchase Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>111233</td>
<td>08/19/2018</td>
<td>1248657</td>
<td>65</td>
<td>HMH-DLU-JUM-DRE</td>
<td>550.00</td>
<td>450.00</td>
<td>100.00</td>
<td>USD</td>
<td>Call center</td>
</tr>
<tr>
<td>2</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>2</td>
<td>HMH-DLU-UIU-FRE</td>
<td>250.00</td>
<td>200.00</td>
<td>50.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>3</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>29</td>
<td>HMH-POU-AWE-SDE</td>
<td>15000.00</td>
<td>13000.00</td>
<td>2000.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>4</td>
<td>111235</td>
<td>11/15/2018</td>
<td>1234568</td>
<td>199</td>
<td>HMH-NHU-DFD-XDS</td>
<td>2500.00</td>
<td>2000.00</td>
<td>500.00</td>
<td>USD</td>
<td>Store</td>
</tr>
<tr>
<td>5</td>
<td>111236</td>
<td>11/17/2018</td>
<td>1234569</td>
<td>43</td>
<td>QEW-GFT-SAD-DEW</td>
<td>947.00</td>
<td>850.00</td>
<td>97.00</td>
<td>USD</td>
<td>Store</td>
</tr>
</tbody>
</table>

- You can imagine that each purchase represents an individual line item associated with a single order (Sales Order ID)

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Sales Transaction Data

- Each row in this table represents the purchase of a particular product (Product SKU)

<table>
<thead>
<tr>
<th>ID</th>
<th>Sales Order ID</th>
<th>Purchase Date</th>
<th>Customer ID</th>
<th>Product SKU</th>
<th>Product Category</th>
<th>Sales Revenue</th>
<th>Sales Cost</th>
<th>Profit</th>
<th>Currency</th>
<th>Purchase Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>111233</td>
<td>08/19/2018</td>
<td>1248657</td>
<td>65</td>
<td>HMH-DLU-JUM-DRE</td>
<td>550.00</td>
<td>450.00</td>
<td>100.00</td>
<td>USD</td>
<td>Call center</td>
</tr>
<tr>
<td>2</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>2</td>
<td>HMH-DLU-UIU-FRE</td>
<td>250.00</td>
<td>200.00</td>
<td>50.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>3</td>
<td>111234</td>
<td>09/22/2018</td>
<td>1248653</td>
<td>29</td>
<td>HMH-POU-AWE-SDE</td>
<td>15000.00</td>
<td>13000.00</td>
<td>2000.00</td>
<td>USD</td>
<td>Online</td>
</tr>
<tr>
<td>4</td>
<td>111235</td>
<td>11/15/2018</td>
<td>1234568</td>
<td>199</td>
<td>HMH-NHU-DFD-XDS</td>
<td>2500.00</td>
<td>2000.00</td>
<td>500.00</td>
<td>USD</td>
<td>Store</td>
</tr>
<tr>
<td>5</td>
<td>111236</td>
<td>11/17/2018</td>
<td>1234569</td>
<td>43</td>
<td>QEW-GFT-SAD-DEW</td>
<td>947.00</td>
<td>850.00</td>
<td>97.00</td>
<td>USD</td>
<td>Store</td>
</tr>
</tbody>
</table>

*Only a sample of data is represented. Awarded teams will have access to the full dataset.*
Total Number of Sales Transactions

- This shows the total number of individual sales transactions (line items) per customer

Takeaway: Customers in a B2B setting make a lot of purchases!

- The histogram to the left represents 2 years of data for 98% of ALL customers
- The remaining 2% made an extremely high number of purchases
Total Number of Sales Transactions

- This shows the total number of individual sales transactions (line items) per customer.

Takeaway: Customers in a B2B setting make a lot of purchases!

- The histogram to the left represents 2 years of data for 98% of all customers.
- The remaining 2% made an extremely high number of purchases.
Total Amount of Revenue

• This shows the total amount of sales revenue per customer

Takeaway: Customers in a B2B setting generate a lot of revenue!

• The histogram to the left represents 2 years of data for 98% of ALL customers
• The remaining 2% generated an extremely high amount of revenue
Total Amount of Revenue

- This shows the total amount of sales revenue per customer

Takeaway: Customers in a B2B setting generate a lot of revenue!

- The histogram to the left represents 2 years of data for 98% of ALL customers
- The remaining 2% generated an extremely high amount of revenue
Questions?

To ask a question, click the chat button in the upper-right and type your question

• All questions will be answered live or via e-mail
Data Structure

Registered Customers

- Sales Transaction Data
- Web Traffic Data
- Email Marketing Data
Web Traffic Data

- Each row in this table shows the *number of daily page views for a registered customer* (Customer ID) who has made a purchase.

<table>
<thead>
<tr>
<th>ID</th>
<th>Date</th>
<th>Customer ID</th>
<th># page views for that day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/19/2018</td>
<td>1248653</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>08/19/2018</td>
<td>1248653</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>08/19/2018</td>
<td>1248653</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>10/09/2018</td>
<td>1248654</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>11/21/2018</td>
<td>1248655</td>
<td>9</td>
</tr>
</tbody>
</table>

- We should note that the details (e.g. URL, search parameters, link type, etc.) for the *individual page views* is available as part of this dataset.

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Data Structure

Registered Customers

- Sales Transaction Data
- Web Traffic Data
- Email Marketing Data
Email Marketing Data

- Each row is a record of an *email campaign event* (e.g. email sent, email opened, etc.) *for an e-mail address associated with a registered customer*

<table>
<thead>
<tr>
<th>ID</th>
<th>Event Captured Date/Time</th>
<th>Customer ID</th>
<th>Hashed Email</th>
<th>Email Permission Status (Opt-In/Opt-Out)</th>
<th>Campaign ID</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/19/2018 9:04:27</td>
<td>570525922</td>
<td>2ec00a41a397dd35e01e99e4bc1298a</td>
<td>I</td>
<td>14565</td>
<td>1 (sent)</td>
</tr>
<tr>
<td>2</td>
<td>08/19/2018 9:05:27</td>
<td>570525923</td>
<td>40a8e364a90b2449f64c7ca9c3309aab</td>
<td>I</td>
<td>14566</td>
<td>4 (opened)</td>
</tr>
<tr>
<td>3</td>
<td>08/22/2018 9:16:27</td>
<td>570525924</td>
<td>5eff2bdc6fcb55648acd8af2f24b179a</td>
<td>I</td>
<td>14567</td>
<td>2 (bounced)</td>
</tr>
<tr>
<td>4</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>a2288ca9898f5f21d72e34570bca05e8</td>
<td>I</td>
<td>14567</td>
<td>14 (viewed_form)</td>
</tr>
<tr>
<td>5</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>fecebff9855b31b0ed19c027782029a7</td>
<td>I</td>
<td>14567</td>
<td>15 (submitted_form)</td>
</tr>
</tbody>
</table>

- The “submitted_form” event means that the customer either *joined the mailing list, requested more information from the company, or enrolled in digital proof of delivery receipts*

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Email Marketing Data

- Each row is a record of an email campaign event (e.g. email sent, email opened, etc.) for an e-mail address associated with a registered customer

<table>
<thead>
<tr>
<th>ID</th>
<th>Event Captured Date/Time</th>
<th>Customer ID</th>
<th>Hashed Email</th>
<th>Email Permission Status (Opt-In/Opt-Out)</th>
<th>Campaign ID</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/19/2018 9:04:27</td>
<td>570525922</td>
<td>2ce00a41a397dd35e01e99e4bc1298a</td>
<td>I</td>
<td>14565</td>
<td>1 (sent)</td>
</tr>
<tr>
<td>2</td>
<td>08/19/2018 9:05:27</td>
<td>570525923</td>
<td>40a8e364a90b2449f64c7ca9c3309aab</td>
<td>I</td>
<td>14566</td>
<td>4 (opened)</td>
</tr>
<tr>
<td>3</td>
<td>08/22/2018 9:16:27</td>
<td>570525924</td>
<td>5eff2bdc6fcf55648acd8af2f24b179a</td>
<td>I</td>
<td>14567</td>
<td>2 (bounced)</td>
</tr>
<tr>
<td>4</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>a2288ca9898f5f21d72e34570bca05e8</td>
<td>I</td>
<td>14567</td>
<td>14 (viewed_form)</td>
</tr>
<tr>
<td>5</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>fecebffe9855b31b0ed19c027782029a7</td>
<td>I</td>
<td>14567</td>
<td>15 (submitted_form)</td>
</tr>
</tbody>
</table>

- Email campaigns typically include things like general promotions, seasonal promotions, product promotions, or conference information

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Email Marketing Data

- Each row is a record of an email campaign event (e.g. email sent, email opened, etc.) for an e-mail address associated with a registered customer

<table>
<thead>
<tr>
<th>ID</th>
<th>Event Captured Date/Time</th>
<th>Customer ID</th>
<th>Hashed Email</th>
<th>Email Permission Status (Opt-In/Opt-Out)</th>
<th>Campaign ID</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/19/2018 9:04:27</td>
<td>570525922</td>
<td>2cec00a41a397dd35e01e99e4bc1298a</td>
<td>I</td>
<td>14565</td>
<td>1 (sent)</td>
</tr>
<tr>
<td>2</td>
<td>08/19/2018 9:05:27</td>
<td>570525923</td>
<td>40a8e364a90b2449f64c7ca9c3309aab</td>
<td>I</td>
<td>14566</td>
<td>4 (opened)</td>
</tr>
<tr>
<td>3</td>
<td>08/22/2018 9:16:27</td>
<td>570525924</td>
<td>5eff2bdc6fcb55648acd8af2f24b179a</td>
<td>I</td>
<td>14567</td>
<td>2 (bounced)</td>
</tr>
<tr>
<td>4</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>a2288ca9898f5f21d72e34570bca05e8</td>
<td>I</td>
<td>14567</td>
<td>14 (viewed_form)</td>
</tr>
<tr>
<td>5</td>
<td>09/19/2018 10:04:27</td>
<td>570525925</td>
<td>fecebbf9855b31b0ed19c027782029a7</td>
<td>I</td>
<td>14567</td>
<td>15 (submitted_form)</td>
</tr>
</tbody>
</table>

- And we should note that there can be multiple e-mail addresses associated with a single customer (Customer ID)

* Only a sample of data is represented. Awarded teams will have access to the full dataset.
Sales Revenue vs. Number of Website Visits

- Each mark represents the total amount of sales revenue and the total number of website visits, for a single customer.

With Outliers

The marks that represent the outliers, which are the customers with a relatively high number of website visits.
Sales Revenue vs. Number of Website Visits

- Each mark represents the total amount of sales revenue and the total number of website visits, for a single customer.

**With Outliers**

The marks that represent the outliers, which are the customers with a relatively high number of website visits.

**Without Outliers – A More “Normal” Distribution**

Removing the outliers, we can see a more “normal” distribution of website visits.

**Takeaway: Do customers who visit the website more often, generate more revenue?**

- The plots suggest that *they do, but only to a point* ...
- Question: *What kind of customers are we observing* visiting the website?
- Question: *And how do you drive new customers to the website?*
Sales Revenue vs. Number of Emails Sent

- Each point represents the total amount of sales revenue and the total number of emails sent, for a single customer.

The points that represent the outliers, which are the customers who received a relatively high number of emails.

Takeaway: Do customers who receive more emails, generate more revenue?

- The plot suggests that they do
- Question: What kind of customers are we observing receiving these emails?
- Question: And are these customers purchasing more often or just purchasing more expensive products?
Questions?

To ask a question, click the chat button in the upper-right and type your question

• All questions will be answered live or via e-mail
Researcher Proposal & Application Process
Research Proposal Format

Researchers will apply and submit a proposal online. Proposals should be in PDF format, no more than 2000 words, and include the following information:

• Title

• Author(s) name, title, affiliation and e-mail address: Please designate a corresponding author (Note: Teams are strongly encouraged, e.g. doctoral student(s) + faculty)

• Author bios: Include up to a paragraph-long biography highlighting what each team member will contribute to the project

• Abstract

• Summary information: Include a single “slide” that visually summarizes the team & project
Research Proposal Format (cont.)

• Introduction: Describe expected contribution(s), covering both the academic and practical aspects. Describe how you will approach the project and the key methods and ideas that you would like to bring to the table. Please keep it concise, and cite relevant work as necessary to explain your academic contribution. There is no need to include a lengthy literature review.

• Detailed project proposal: Please include supporting detail that will help us assess the feasibility of your approach and its compatibility with existing data.

• Data Needs: Bulleted list of data required or requested for analysis. While we can’t guarantee the inclusion of these items, we are happy to investigate the availability.

• Languages/tools: What you propose to use and how it fits the project. (Note: There are no restrictions on software)
How to Apply

Submit your proposal here: [https://wcai.wharton.upenn.edu/research/b2b-collaborative-ro/](https://wcai.wharton.upenn.edu/research/b2b-collaborative-ro/)

• No later than Monday, February 4th, 2019

In addition to your proposal, we will be organizing some initial (virtual) pre-selection meetings with the company

• To discuss the project opportunity
• To determine if there is a good match in interests

Please contact us at wcai-research@wharton.upenn.edu if you have any questions prior to submitting your proposal
Research Proposal Selection Process

Two research teams will be selected for the Research Opportunities

• Teams will meet regularly with the sponsor, providing opportunities to share knowledge, clarify the business context, and refine the data
• While both teams will be working on CLV, they will be methodologically distinct

Timeline

• Within 6 weeks, several research teams will be selected for initial (virtual) meetings with the company
• Within 9 weeks, two final research teams will be selected & notified
• Within 12 weeks, the projects will kick off
Research Proposal Selection Process (cont.)

Proposals will be evaluated based on:

• Novel approach to the project problem(s)
• Academic contribution and potential to significantly improve the research sponsor’s marketing practice
• Willingness to commit to a highly collaborative project, including regular interactions
• Willingness to share all code and findings with the company
• History of academic achievement
Proposal Review Committee

Research teams will be selected jointly by a committee of academics and company representatives:

- Eric Bradlow: Faculty Co-Director, WCAI
- Raghu Iyengar: Faculty Co-Director, WCAI
- Elea Feit: Senior Fellow, WCAI
- Data scientists from the company
- Other external reviewers
Other WCAI Opportunities for Researchers

If you registered for this webinar, you will receive regular announcements regarding upcoming Research Opportunities:

• Research Projects: http://wcai.wharton.upenn.edu/for-researchers/research-opportunities/

• Research Papers: http://wcai.wharton.upenn.edu/for-researchers/research-paper-series/

• Sign up for updates: http://wcai.wharton.upenn.edu/newsletter/

• E-mail us: wcai-research@wharton.upenn.edu
Questions?

To ask a question, click the chat button in the upper-right and type your question
• All questions will be answered live or via e-mail