Text4baby Data Analysis Toolkit for Partners

Introduction

Text4baby is a data-driven initiative. Staff routinely monitor a wide array of data collected by the program to inform and improve promotional, outreach and product strategies. We believe that data-driven decision-making is critical to guide and target text4baby outreach efforts effectively and efficiently. We provide our partners with the data and training to support them in identifying efforts that drive enrollment in their own communities. Partners who have used text4baby data to plan, implement and evaluate their outreach efforts have experienced great success in driving enrollment and reaching mom’s with critical health information!

This toolkit is designed for our partners as a resource to utilize and analyze data from the text4baby portal. The toolkit is organized by questions that outreach partners may have, which can be answered using the data available in the portal. In each section we provide a brief overview of how to answer each question using text4baby data and also link to additional resources, including webinars and step-by-step guides located in the appendix. The toolkit also includes a section on additional ideas for evaluating text4baby in your community.

To Get Started

(1) Complete a Data Use Agreement: Once we receive your completed and signed Data Use Agreement via fax or email, we will work with our technology vendor to create log-in credentials to give you access to real-time, individual-level data for your state.

(2) Review Instructions & Data Definitions (Appendices H & I): Along with your credentials, you will receive two documents that will provide you an overview of how to use the text4baby portal and definition of the data available on the portal.

(3) Monthly Webinars: We will host monthly webinars to assist you in accessing and using the text4baby data. Partners will receive monthly emails with instructions on how to sign up for the webinar if interested.

Before you start, make sure you know how to view, export, and filter data in the text4baby portal and that your login credentials are up to date.
Introduction .................................................................................................................. 1

To Get Started ............................................................................................................. 1

1) How can I find out the percentage of users who are currently subscribed and receiving messages? ................................................................................................................. 1

   a. Counts of Subscribed vs. Unsubscribed Users .......................................................... 1
   b. Counts of Users Receiving Protocol vs. Ad Hoc Messages ............................................. 2

2) How can I tell if my outreach activity was effective? .................................................... 3

   a. Time Series Graphs by Date ....................................................................................... 3
   b. Time Series Graphs by Time of Day ............................................................................. 4

3) How can I use the characteristics of my community’s current text4baby users to plan future outreach initiatives in my community? ........................................................................ 4

   a. Counts of English vs. Spanish Users ........................................................................... 4
   b. Counts of Users by Type ............................................................................................... 6

4) How can I decide which zip codes or counties I should target for outreach? .............. 7

   a. Compare text4baby users by zip code/county to total estimated pregnant women/new moms by zip code/county ..................................................................................... 7

5) How do I know if people are using my organization’s text4baby web button? .......... 8

   a. Referral URL tab ......................................................................................................... 8

6) Evaluation Ideas ....................................................................................................... 9

7) What if I have more questions? ................................................................................ 10

Appendix ....................................................................................................................... 11

Appendix A: Counting Subscribed vs. Unsubscribed Users ......................................... 11

Appendix B: Counting Users Currently Receiving Weekly Protocol Messages ............ 12

Appendix C: Time Series by Date .................................................................................. 13

Appendix D: Time Series by Time of Day ....................................................................... 17

Appendix E: Counts of English vs. Spanish Users, by county ...................................... 21

Appendix G: Analyzing Users and Live Births by Zip Code .......................................... 26

Appendix H: Accessing Text4baby Data ..................................................................... 29

Appendix I: Outreach Partner Data Definitions ............................................................. 32
1) How can I find out the percentage of users who are currently subscribed and receiving messages?

a. Counts of Subscribed vs. Unsubscribed Users

Comparing subscribed users of text4baby to those who have unsubscribed will allow you to understand the number of users in your community currently receiving text4baby messages (weekly protocol messages and/or alert/adhoc messages). It will also allow you to see how unsubscription in your community compares to the national average available on the dashboard tab of the portal.

People may unsubscribe for a variety of reasons, such as completion of the weekly text4baby messages (and subsequent opting out of receiving alert/adhoc messages), being outside of the target audience, or a misunderstanding of what the text4baby program actually entails. Not every user who completes the new baby protocol (baby is older than one year) unsubscribes from text4baby service and is thus still eligible to receive alert/adhoc messages.

Example: In the State of Texas, I am interested understanding how many users are currently receiving messages and seeing how the percentage of unsubscribed users compares to the national percentage.

![Users by Subscription Status: Texas](image)

Over 58% (roughly 22,200) of the users who have ever enrolled in text4baby from Texas are currently receiving protocol and/or adhoc/alert messages (see section below to calculate how many users are receiving...
weekly protocol vs adhoc/alert messages). The percentage of users in Texas receiving messages is similar to percentage (58%) of subscribed users across the county (available on dashboard tab).

For Step-by-Step guides on how to conduct counts of subscribed and unsubscribed users, please see the appendix, and refer to the online video guide.

b. Counts of Users Receiving Protocol vs. Ad Hoc Messages

Comparing the number of users by message type (protocol vs. ad hoc) among subscribed users will allow you to understand the number of users in your community currently receiving weekly protocol messages vs the number only receiving adhoc/alert messages. Not every user who completes the new baby protocol (baby is older than one year) unsubscribes from text4baby service and is thus still eligible to receive adhoc/alert messages.

Example: I want to see the percentage of users who are currently receiving weekly protocol messages in Texas.

![Pie chart showing the percentage of users receiving protocol messages vs ad hoc messages.](chart.png)

The majority of users (77%) are currently receiving either pregnancy or new baby weekly messages. The remaining 23% represent new moms with baby’s >1 who are no longer receiving weekly messages and are only receiving adhoc/alert messages.
For Step-by-Step guides on how to conduct counts of users currently receiving protocol messages, please see the appendix.

2) How can I tell if my outreach activity was effective?

a. Time Series Graphs by Date

If you know the start (and/or end) dates of your outreach activity, a time series graph by date can tell you if enrollment was higher than usual during that time period.

**Example:** Outreach partners in North Carolina began running their own text4baby PSA on a few major networks beginning in July of 2012. In order to see if this was effective at driving enrollment we created the following time series.

**NOTE:** This chart could also be complete using just text4baby data; additional data (in this case, # of PSA airings) are not necessary to determine if enrollment was higher than usual during an outreach activity.

We can see in the above graph that the weekly registration peaks became higher beginning in July, when the PSA started airing. Closer analysis revealed that 1,203 users registered during July, compared to only 597 during June. This is an increase of over 200%.

For Step-by-Step guides on how to conduct time series, please see the appendix, and refer to the online video guide.
b. Time Series Graphs by Time of Day

Once a time series by date has been conducted, you may want to look at certain dates in detail to determine if registration was concentrated around a certain time of day. For example, if you were running PSA’s you may want to see if registration matches the time of day the PSA aired, or if you had ads on public transit, you may want to see if there are registration spikes during rush hour.

Example: Using the time series by date graph on page 1, we saw an registration spike on July 26. We want to know if North Carolina’s PSAs drove this spike, so we conducted the following time series based on time of day. We know that on July 26, PSAs aired at 10:15 AM, 1:28 PM, 1:58 PM, 2:35 PM, 3:26 PM, 3:59 PM, 4:32 PM, and 5:00 PM.

We can see in the above graph that registrations on July 26 spiked around 12:38 PM and 2:40 PM. This suggests that the 2:35 PM PSA is likely responsible for the large spike that began at 2:35 PM and ended at 2:41 PM.

For Step-by-Step guides on how to conduct time series, please see the appendix, and refer to the online video guide.

3) How can I use the characteristics of my community’s current text4baby users to plan future outreach initiatives in my community?

a. Counts of English vs. Spanish Users
Comparing the percentage of text4baby users in the Spanish protocol to population-level data on the percentage of Spanish speakers in your community may help you target outreach to Spanish speakers in selected counties.

**Example:** We know that Illinois has a large population of Hispanics, but we do not know if our outreach is effectively targeting Spanish speakers in selected counties. Using the Census FactFinder, we compared the percentage of text4baby users enrolled in the Spanish protocol to the percent of people who speak Spanish in the home for select counties in Illinois.

<table>
<thead>
<tr>
<th>County</th>
<th>Number t4b users</th>
<th>Number t4b users in the Spanish Protocol</th>
<th>% Users in Spanish Protocol</th>
<th>Percent Spanish speaking 2010</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Champaign</td>
<td>123</td>
<td>1</td>
<td>0.81%</td>
<td>4.80%</td>
<td>3.99%</td>
</tr>
<tr>
<td>Cook</td>
<td>9169</td>
<td>1056</td>
<td>11.52%</td>
<td>20.30%</td>
<td>8.78%</td>
</tr>
<tr>
<td>Kane</td>
<td>961</td>
<td>184</td>
<td>19.15%</td>
<td>25.90%</td>
<td>6.75%</td>
</tr>
<tr>
<td>Lake</td>
<td>653</td>
<td>86</td>
<td>13.17%</td>
<td>17.00%</td>
<td>3.83%</td>
</tr>
<tr>
<td>McHenry</td>
<td>214</td>
<td>7</td>
<td>3.27%</td>
<td>9.30%</td>
<td>6.03%</td>
</tr>
<tr>
<td>Peoria</td>
<td>140</td>
<td>1</td>
<td>0.71%</td>
<td>2.80%</td>
<td>2.09%</td>
</tr>
<tr>
<td>Will</td>
<td>631</td>
<td>57</td>
<td>9.03%</td>
<td>11.80%</td>
<td>2.77%</td>
</tr>
<tr>
<td>Winnebago</td>
<td>431</td>
<td>29</td>
<td>6.73%</td>
<td>8.20%</td>
<td>1.47%</td>
</tr>
</tbody>
</table>

This analysis shows us that text4baby users in selected Illinois counties may not fully represent the population in each county. The largest disparity between the percentage of Spanish speakers and the
percentage of Spanish language text4baby users is seen in Cook County, with a nearly 9% difference. This information can be used to target outreach to Spanish-speaking populations in selected counties.

For Step-by-Step guides on how to conduct counts of user types and protocols, please see the appendix, and refer to the online video guide.

b. Counts of Users by Type

Beginning on February 29, 2012, new enrollees of text4baby were required to report their user type (Pregnant, New Mom, Dad/Parent, Friend/Relative, or Health Care Provider/Observer) to enroll in the service. In August 2012, text4baby sent an adhoc message to users who registered prior to February 29, 2012 and were still subscribed, asking them to provide their user type. Data is available for all users who enrolled after 2/29/2012 and for some users who enrolled prior to that date and responded to the adhoc survey.

You can compare the breakdown of users by type to see what types of users are enrolling in your community.

Example: In Illinois, we want to know what percentage of enrollees are in our target audience (Pregnant or New Mom). We generated counts of users in Illinois by type.

This chart reveals that the overwhelming majority (93%) of users in Illinois are in our target groups (Pregnant or New Mom). You can also filter by date of registration (2/29/2012), to calculate the percentage of pregnant of
pregnant women vs new mom who initially register for text4baby. National data suggested a higher percentage of pregnant women initially register, suggesting additional outreach to new moms should be considered.

For Step-by-Step guides on how to conduct counts of user types and protocols, please see the appendix, and refer to the online video guide.

4) **How can I decide which zip codes or counties I should target for outreach?**

a. **Compare text4baby users by zip code/county to total estimated pregnant women/new moms by zip code/county**

Comparing text4baby enrollees to live birth data and/or census data can give you a sense of need in your community and help you see if you are reaching areas with a large segment of the target audience.

Example: We want to know what zip codes to target in San Diego County to reach the greatest number of pregnant women and new moms live in zip codes with the highest poverty. Using California 2011 live birth data, we were able to look at live births on a local level and compare them to the percentage of users who enrolled in 2011 from those zip codes.
<table>
<thead>
<tr>
<th>Highest Poverty Zip Codes</th>
<th>Number of t4b Users in 2011</th>
<th>Percentage of t4b Enrollees in SD County</th>
<th>Estimated Pregnancies and New Moms(=2011 Live Births*2)</th>
<th>Percentage of Estimated Pregnancies and New Moms</th>
</tr>
</thead>
<tbody>
<tr>
<td>92105</td>
<td>58</td>
<td>3.4%</td>
<td>2376</td>
<td>2.7%</td>
</tr>
<tr>
<td>92113</td>
<td>50</td>
<td>3.0%</td>
<td>2152</td>
<td>2.5%</td>
</tr>
<tr>
<td>92173</td>
<td>24</td>
<td>1.4%</td>
<td>1670</td>
<td>1.9%</td>
</tr>
<tr>
<td>92115</td>
<td>36</td>
<td>2.1%</td>
<td>1774</td>
<td>2.0%</td>
</tr>
<tr>
<td>92102</td>
<td>35</td>
<td>2.1%</td>
<td>1588</td>
<td>1.8%</td>
</tr>
<tr>
<td><strong>Total Users in SD county 2011 = 1691</strong></td>
<td></td>
<td></td>
<td><strong>Total estimated pregnant women and new moms in SD county in 2011 = 87,902</strong></td>
<td></td>
</tr>
</tbody>
</table>

This analysis tells us the proportion of text4baby users in most of the highest poverty zip codes is greater than the proportion of estimated pregnancies and live births in those zip codes. Only one in one zip code – 92183 – do we see a lower proportion of text4baby users as compared to the proportion of estimated pregnancies and new moms. Zip code 92173 represents 1.9% of estimated San Diego county pregnant women and new moms in San Diego county, but only 1.4% of text4baby enrollees are from that zip code. Therefore the 92173 zip code should be considered for targeted text4baby outreach.

The CDC Natality data site is a good source for live birth data by county. This data source can show live births by county, race/ethnicity, mother’s education level, marital status and age, as well as birth characteristics and risk factors. Data is not available for counties with fewer than 100,000 people. Your state health department may have information on live births in smaller counties, or by zip code. Check out their website to see.

For a step-by-step guide on calculating total text4baby users by county or zip code, please refer to the appendix or to the online video guide.

5) **How do I know if people are using my organization’s text4baby web button?**

a. **Referral URL tab**

If your organization has placed a text4baby web button on your web page, you can use the portal to see if people are using it to sign up for text4baby.
Example: MTV has placed a text4baby web button on their page for *16 and Pregnant* and wants to know if the button is effective. Using data from the portal, they found that 576 people have used their web button to register, accounting for 3% of all web signups in the country.

### 6) Evaluation Ideas

In addition to analyzing text4baby data, other evaluation activities can be conducted to learn more about the impact of text4baby in your community. Some evaluation ideas are included below:

**a. Focus/Discussion Groups:** Several text4baby outreach partners have conducted focus/discussion groups with text4baby users or women who represent the text4baby target population to learn more about user receptiveness of text4baby. Partners have tested our outreach materials, explored topics of interest to help inform text4baby messaging, and learned more about the features of the service liked most by users through focus/discussion group efforts.

Other partners have held focus/discussion groups with other text4baby stakeholders, such as healthcare providers, state/local health departments, and Medicaid/CHIP and WIC program staff, to learn more about the adoption of the service, among other topics of interest, from stakeholders who are interacting directly with text4baby users.

**b. Partner with an Academic or Research Organization to Conduct a Formal Evaluation:** Five academic institutions and one research organization have conducted or are currently conducting evaluations of text4baby. Text4baby outreach partners have partnered with these organizations to work together to learn more about attitude, knowledge and behavior change among text4baby users. Check out the results of the first randomized evaluation of text4baby. Consider partnering with an academic or research organization to explore how you can learn more about text4baby in your community!

**c. Evaluate Text4baby Using Your Electronic Health Record:** One text4baby outreach partner recently added structured data fields within their network-
wide EHR that capture whether a patient is enrolled in text4baby. The fact that text4baby enrollment data are in the same electronic system as prenatal and birth outcome data positions an organization to assess the impact of text4baby on prenatal care and birth outcomes.

7) **What if I have more questions?**

Please contact Jessica Bushar, text4baby Research Manager at jbushar@hmhb.org with additional questions or comments.
Appendix

Appendix A: Counting Subscribed vs. Unsubscribed Users

1. Pull your data from the portal.
2. Generate counts of subscribed and unsubscribed users.
   a. To the right of your large data table, create a small table with two rows, labeled “On” and “Off”.
   b. Input the following formula in the cell to the right of the “On” label:
      \[ \text{COUNTIF}(C:C, \text{M3}) \]
      i. Scroll your mouse cursor over the bottom right corner of each cell until it forms a little black plus sign, then drag to the row labeled “Off”.
3. Create and format a pie chart.
   a. Copy the rows labeled On and Off and paste onto a new tab in the Excel document. Be sure to paste “values only”
   b. Highlight the data table and select Pie Graph from the Insert tab in Excel.
   c. Use the “Layout” tab in “Chart tools” to label data and title the graph.
Appendix B: Counting Users Currently Receiving Weekly Protocol Messages

1. Pull your data from the portal.
2. Filter data by SMS subscription = “on”

3. Generate counts of currently subscribed users who are currently are and are not receiving weekly protocol messages.
   a. To the right of your large data table, create a small table with two rows, labeled “Yes” and “No”.
   b. Input the following formula in the cell to the right of the “Yes” label: 
      \[ \text{COUNTIF}(D:D, M3) \]

4. Generate counts of new baby users currently receiving protocol messages to determine the number of users currently receiving protocol vs. ad hoc messages in the new baby protocol.
   a. To the right of your large data table, create a second small table with one row labeled ‘new baby”
   b. Input the following formula in the cell to the right of the “Yes” label: 
      \[ \text{COUNTIF}(D:D, M3, B:B, O3) \] This will give you the number of new baby users currently receiving weekly protocol messages. Repeat the same
steps for “pregnancy” to see how many users are currently receiving weekly pregnancy protocol messages.

5. Create and format a pie chart.
   a. Copy the rows labeled Yes and No and paste onto a new tab in the Excel document. Be sure to paste “values only”
   b. Highlight the data table and select Pie Graph from the Insert tab in Excel.

Appendix C: Time Series by Date
1. Begin by exporting your state’s data from the participant tab on the portal. You may also narrow it by county or zip code.
   a. You may want to filter by date, so that you only pull new users who enrolled during your outreach activity, and a month or so before or after.
2. Reformat the Registration Date stamp on your Excel document
a. Insert four columns next to the column labeled Registration Date.

<table>
<thead>
<tr>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Registration</td>
<td>Date of Regi...</td>
<td>DAY</td>
<td>MONTH</td>
<td>YEAR</td>
</tr>
<tr>
<td>Mar 01, 2013 10:10 AM</td>
<td></td>
<td>=DAY(H2)</td>
<td>=MONTH(H2)</td>
<td>=YEAR(H2)</td>
</tr>
<tr>
<td>Mar 01, 2013 10:03 AM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. Label the first column to the right “Day” and input the following formula in the first data field: =DAY(H2) . The number should match the day of the month in the registration date column.

c. Label the next column “Month” and input the following formula: =MONTH(H2)

d. Label the next column “Year” and input the following formula: =YEAR(H2)

e. In the final column, label Date and input the following formula: =DATE(K2,J2,I2) . This will give you the registration date, with the registration time removed.

<table>
<thead>
<tr>
<th>J</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel of Regi...</td>
<td>Registration Date</td>
<td>DAY</td>
<td>MONTH</td>
<td>YEAR</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mar 01, 2013 10:10 AM</td>
<td>=DAY(H2)</td>
<td>=MONTH(H2)</td>
<td>=YEAR(H2)</td>
<td>=DATE(K2,J2,I2)</td>
<td></td>
</tr>
</tbody>
</table>

i. Scroll your mouse cursor over the bottom right corner of each cell until it forms a little black plus sign, then double click. This will apply the formula to the other cells in the column.

3. Format the data for your table.
a. Copy the new “Date” column. Paste it onto a new tab in your workbook, making sure to paste values only.

![Image of Excel worksheet showing a copied column]

b. Paste one more copy of the Date column to the right and relabel it “Date Deduped”.

c. Highlight the “Date Deduped” column and select “remove duplicates” from the “Data” tab in Excel.

![Image of Excel worksheet showing remove duplicates dialog]

Next to the “Date Deduped” column, form a new column labeled “Number of Users”. Input the following formula in the first cell of the column: =COUNTIF(A:A, C2). This formula will tell you how many entries in the original date column equal the date in your “Date Deduped” column.

![Image of Excel worksheet showing the “Number of Users” column]

i. Scroll your mouse over each of these cells until it forms a little black plus sign, then double click on the bottom right corner. This will apply the formula to every other cell in the column.
4. Create and format a line graph.
   a. Highlight the entries in the “Date Deduped” and “Number of Users” columns, and select a line graph from the Insert tab in Excel. This will show how the number of new enrollees change over time and if there any specific days or time periods where enrollment spiked.

   b. Use the “Layout” tab in “Chart tools” to label axes and title the graph.
Appendix D: Time Series by Time of Day

1. Begin by using the same Excel document that you used to make a time series based on date.
2. Reformat the time stamp in the Registration Date column in the Excel file that you used to make your time series based on date.
   a. Insert four columns next to the column labeled Registration Date.
      b. Label the first column to the right “Second” and input the following formula in the data first empty cell: \( =\text{SECOND}(H2) \). In this document, every “Second” value will be zero.
   c. Label the next column “Minute” and input the following formula: \( =\text{MINUTE}(H2) \)
   d. Label the next column “Hour” and input the following formula: \( =\text{HOUR}(H2) \)
   e. In the final empty column, label “Time” and input the following formula: \( =\text{TIME}(K2,J2,I2) \). This will give you the time of day of enrollment, with the date removed.
   i. Scroll your mouse cursor over the bottom right corner of each cell until it forms a little black plus sign, then double click. This will apply the formula to the other cells in the column.
4. Format the data for your table.
   a. Copy the new “Time” column. Paste it onto a new tab in your workbook, making sure to paste values only. Relabel this column “Time Deduped” (You may need to reformat the cells to show time after you have pasted)
   
   ![Image of Excel paste options]

   b. Highlight the “Time Deduped” column and select “remove duplicates” from the “Data” tab in Excel.

   ![Image of Excel data tools]

   c. With the “Time Deduped” column highlighted, select Sort from the Data tab in Excel, and sort by smallest to largest.

   ![Image of Excel sort options]
6. Select Data Points for Table.
   a. Using the Filter function in the reformatted “Date” column, show only the new users for a specific date.
   b. In the filtered view, copy the reformatted “Time” column, and paste it into the other tab, a few columns to the right of the “Time Deduped” column. Make sure to paste as “values only” and format as necessary so it shows time.

7. Generate counts of new users.
   a. Create a new column labeled “Number of Users” next to the “Time Deduped” column.
   b. Input the following formula in the first cell of the column:
      \[=\text{COUNTIF}(D:D,A2)\]. This formula will tell you how many entries in the “Time” column for your date of interest match the time in the “Time Deduped” column.
8. Create and format a line graph.
   a. Highlight the entries in the “Time Deduped” and “Number of Users” columns, and select a line graph from the Insert tab in Excel. This will show how the number of new enrollees change over the course of the day and if there any specific time periods where enrollment spiked.

   ![Line Graph Example]

   b. Use the “Layout” tab in “Chart tools” to label axes and title the graph.
Appendix E: Counts of English vs. Spanish Users, by county

1) Begin by exporting your state’s data from the participant tab on the portal.
2) Generate counts of Spanish and total users by county.
   a) To the far right of your dataset, list out the counties you are interested in analyzing. To the right of your new “County” column, create two new columns title “Number of Users” and “Number of Spanish Users”.
   b) Input the following formula into the “Number of Users” column: 
   \[ \text{COUNTIF}(J:J, M2) \]
   c) Input the following formula into the “Number of Spanish Users” column: 
   \[ \text{COUNTIFS}(J:J, M2, E:E, "Spanish") \]
   i) Scroll your mouse cursor over the bottom right corner of the top cell in the Number of Users and Number of Spanish Users column until it forms a little black plus sign, then drag to the bottom of the column.
3) Compare percentages of Spanish users to population data.
a) Copy and paste the table of user counts by county onto a new tab in your Excel document. Make sure to paste values only.

b) To the right of the table, create a new column labeled “Percent Spanish Users”.

c) Input the following formula in the next empty cell of the Percent Spanish Users column: \( =\frac{C2}{B2} \)

i) Scroll your mouse cursor over the bottom right corner of the top cell in the Number of Users and Number of Spanish Users column until it forms a little black plus sign, then drag to the bottom of the column.

ii) Format cells so that they are seen as percentages, instead of decimals.

d) Using your source of population-level data, input -in a new column- the percentage of Spanish speakers in each county.

e) Copy and paste (values only) the County, Percent Spanish T4B users, and Percent Spanish speaking columns to a new spot on the page.

4) Create and format a bar graph.
a) Highlight the entries in the County, Percent Spanish T4B users, and Percent Spanish speaking columns, and select the clustered column graph from the Insert tab in Excel. This will create a graph comparing the percentages of Spanish text4baby users to the percent of Spanish speakers in each county.

b) Use the “Layout” tab in “Chart tools” to label axes and title the graph.

Appendix F: Counts of Users by Type

1. Export data from the portal for the state (or county or zip code) and date range you are interested in. (Note: User type is only available for all users who enrolled after February 29, 2012. Some users who enrolled prior to February 29, 2012 may have provided their user type when asked via an adhoc survey).

2. Format Data table.
   a. Copy the “Type” column, and paste onto a new tab in your Excel document.
   b. Copy and paste the “Type” column once more onto the worksheet, and relabel it “Type Deduped”.

23
c. Remove duplicates from the “Type Deduped” column by highlighting the column, and clicking on “Remove duplicates” from the Data tab in Excel.

d. Label the column next to “Type Deduped” as “Number of Users”.

e. Input the following formula into the “Number of Users” column:

\[ =\text{COUNTIF}(A:A, \ C2) \]
4. Create and format a pie chart.
   a. Highlight the Type Deduped and Number of Users columns and select the pie graph from the Insert tab in Excel.
   b. Use the “Layout” tab in “Chart tools” to label axes and title the graph.
Appendix G: Analyzing Users and Live Births by Zip Code

1. Export relevant data from the portal. You may want to filter by county or registration date before you export the data.
2. Generate counts of users by zip code (note: you can use this same process to count by county).
   a. Copy the Zip Code column and paste onto a new worksheet.
   b. Copy and paste the Zip Code column one more time and relabel as “Zip Code deduped.” Click “Remove Duplicates” on the Data tab in Excel.
   c. Next to the Zip Code Deduped column, create a new column titled “Number of Users”. In the first cell of the column, input the following formula: =COUNTIF(A:A, C2)
      i. Scroll your mouse cursor over the bottom right corner of each cells until it forms a little black plus sign, then drag to the bottom of the column. This will apply the formula to every other cell in the column.
3. Compare the number of live births by county or zip code to the number of enrollees listed.
   a. Next to the “Number of Users” column, label a new column “Estimated Pregnancies and New Moms”. Use your CDC or local natality data to fill in the number of live births for each zip code or county. Multiply the number of live births by 2 to get estimate pregnancies and new moms.
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zip Code</td>
<td>Zip Code Deduped</td>
<td>Number of Users</td>
<td>Estimated Pregnancies and New Moms</td>
<td>Percentage of Users</td>
</tr>
<tr>
<td>92105</td>
<td>92105</td>
<td>58</td>
<td>2376</td>
<td></td>
</tr>
<tr>
<td>92173</td>
<td>92173</td>
<td>24</td>
<td>1670</td>
<td></td>
</tr>
<tr>
<td>92026</td>
<td>92102</td>
<td>35</td>
<td>1588</td>
<td></td>
</tr>
<tr>
<td>92102</td>
<td>92115</td>
<td>36</td>
<td>1774</td>
<td></td>
</tr>
<tr>
<td>92105</td>
<td>Total</td>
<td>50</td>
<td>2152</td>
<td></td>
</tr>
<tr>
<td>92113</td>
<td>Total</td>
<td>1691</td>
<td>87,902</td>
<td></td>
</tr>
</tbody>
</table>

b. At the bottom of the “Number of Users” column, input the following formula to calculate the total number of users:

=SUM(D2: D95)

c. Input the following formula at the bottom of the “Estimated Pregnancies and New Moms” column to calculate the total number of estimated pregnancies and new moms:

=SUM(E2:E95)

d. Next to the “Estimated Pregnancies and New Moms” column, label a new column “Percentage of Estimated Pregnancies and New Moms”

e. Input the following formula in the “Estimated Pregnancies and New Moms” column:

=(E2/$E$95)*100

i. Scroll your cursor over the formula cell in this column until you see a little black plus sign in the bottom right corner and double click to apply the formula to all the cells.

ii. Format the cells as necessary to show the number as a percentage.

f. Next to the “Percentage of Estimated Pregnancies and New Moms” column, label a new column “Percentage of Users”.

g. Input the following formula in the “Percentage of Enrollees” column:

=(D2/$D$95)*100

i. Scroll your cursor over the formula cell in this column until you see a little black plus sign in the bottom right corner and double click to apply the formula to all the cells.

ii. Format the cells as necessary to show the number as a percentage.

4. Create and format a bar graph.

a. Copy the row of data labels, and the rows containing the zip codes of interest, and paste onto a new tab in Excel. Be sure to paste as values only.

b. Delete the “Number of Users” and “Estimated Pregnancies and New Moms” columns, so only the zip codes and percentages columns are visible.
c. Highlight the data table and select the clustered column graph from the Insert tab in Excel.

d. Use the “Layout” tab in “Chart tools” to label data and title the graph.
Appendix H: Accessing Text4baby Data

Introduction & Overview

To obtain log-in credentials for the text4baby database, your organization must complete the Outreach Partner Data Use Agreement. Please contact Jessica Bushar at jbushear@hmhb.org for more information.

Please save this link to the text4baby portal as a bookmark. We recommend you use Google Chrome to increase the speed of the data portal.

Upon logging in, you will view four tabs across the top of the screen:

1) National Dashboard
2) View Data Outreach Partner
3) Charts
4) Change Password

Overview of the Tabs

The National Dashboard tab will provide you with some key facts about national registration. Please click on the underlined words to view the definitions associated with various terms. You will also find a list of definitions for the items in the Dashboard within the Data Definitions document.

The T4b Dashboard- Outreach Partner tab has a sidebar to the left, which provides access to the following data:

- The Participant’s View for Outreach Partners tab shows individual-level user information for text4baby registration in your state (or nationally if you are a national-level partner). We strongly encourage you to review the Data Definitions document and use it in conjunction with the data on this page.

- The Participant’s by Referral URL tab includes registration data for users who first registered for text4baby via the web; data for those who registered via SMS or API are NOT displayed. This tab is helpful if you are interested in looking at data specific to users, who register via the web, rather than downloading the large set of data available in the Participants for Outreach Partners tab and then filtering the data to focus on web registrations.

- The Referral Survey tab *NEW* includes data for users who were asked a referral survey (first implemented with users who registered on 2/20/2012) to understand the referral sources of users who register in text4baby. This dataset includes records for all users who were asked the referral survey, including those who do not respond to the survey.
The Charts tab includes three charts with national-level data. Please note that these charts are very slow to load.

The Change Password tab allows you to change your password at your convenience.

Using Charts on the National Dashboard

The 2 charts are minimized in size to fit on the page. To expand the chart to its full size, simply click on the image. You can also right-click and choose Save As to download an image of the chart as a JPG.

If you’d like to export the data from these charts, simply click on the Show Data button on the bottom right (do this from the home screen of the Dashboard, not from the expanded version of the chart; it won’t work if you have the expanded version open on your screen). Scroll down and you can click on Export to obtain an Excel spreadsheet of the data presented in the 2 charts.

Export and Filter Data in the Participant’s View for HMHB Outreach Partner Tab

The Outreach Partner tab has a sidebar to the left, which provides access to three datasets. Instructions on how to filter and export datasets are included below:

To download and manipulate the data set, click on EXPORT. You will be able to choose how many pages you would like to download as well as the format. This is the easiest way to analyze data, using the process detailed in the next section below. Make sure to select “All Pages” when you go to export data (the default is Current Page). Please note that it may take 10-15 minutes to download a full data set for your state, given the volume of data. Please be patient. As long as the page is still loading the system is working.

To filter the users in your state by other criteria, click on the FILTER tab and select “QUICK” from the drop-down box. You can view specific data this way, such as only users in the Spanish protocol, only those that are subscribed, or only those in a specific zip code or county. Type in the criteria you want to search on and hit ENTER on your keyboard (there is NOT a button on the screen to execute the search, so you will have to use your keyboard’s ENTER function). You can click on the FILTER tab and select “Date Range” to filter by date of registration if you want to view records of users who registered during a specific time period. You can download filtered data using the same process described above.

Using the Charts Tab

There are three charts included in the Charts Tab which display additional national-level data on text4baby. These include:

- CHART 1 - Users Registered Per Hour for Past 30 Days
• CHART 2 - Users Who Have Registered Per Day, Since Text4baby Launch on 2/2/2010, including Cumulative Total Users Who Have Registered Since Launch
• CHART 5 - Total Users Distributed by Week of Pregnancy Upon Original Registration in Pregnancy Protocol, Since Text4baby Launch on 2/2/2010

If you’d like to export data from these charts, simply click on the Show Data button on the bottom left of any of the charts. Scroll down and you can click on Export to obtain an Excel spreadsheet of the data.
Appendix I: Outreach Partner Data Definitions

Generic Reference Definitions:

User
- An individual who has completed a registration form (SMS, web or API) using a unique phone number.
- There will only be one user in text4baby associated with a unique phone number.
- If a user unsubscribes and re-subscribes in the system using the same phone number, this user is only counted once using the first day that they registered.

Interaction

Users can interact with the text4baby system in four ways:

1) Registration by web, SMS or API (application programming interface)
2) Update by text
3) Unsubscription by text
4) Respond to survey questions by text
Dashboard Data Definitions (Tab 1)

The Dashboard provides an overview of real-time registration data at the national- and state-levels. The elements listed below are found in the tab labeled “National Dashboard”. Please note these data are updated each morning at 12:00am EST.

**Number of New Users**

Displays the number of users who have registered in text4baby by SMS, web or API within the time periods listed (Yesterday; Last 7 days; Last 30 Days; Last 365 Days). “Yesterday” covers a period running from 12:00am EST on current day to 12:00am EST on previous day.

**Number of Messages Sent**

Displays the number of mobile-terminated SMS text messages that have been sent to text4baby users within the time periods listed (Yesterday; Last 7 Days; Last 30 Days; Last 365 Days). “Yesterday” covers a period running from 12:00am EST on current day to 12:00am EST on previous day.

**Fast Facts Box:**

*Total number of users since launch (2/2/2010)*

Cumulative total number of users who have enrolled in text4baby since launch on 2/2/2010.

*Total number of messages sent since launch (2/2/2010)*

Cumulative total number of mobile-terminated SMS text messages that have been sent to text4baby users since launch on 2/2/2010.

**Weekly Growth Rate (Compared to 2012 Average) *NEW***

This weekly growth rate calculation is a useful indicator to look at weekly growth in registrations relative to average weekly registrations in 2012. This growth rate compares total weekly registrations (Sunday – Saturday) to average weekly registrations for 2012. The field is updated every Sunday morning at 12am EST.

\[
\text{Weekly growth rate (compared to 2012 average)} = \frac{(X-Y)}{Y} \times 100
\]

Where \(X\) = Total weekly registrations
\(Y\) = Average weekly registrations in 2012

**Weekly Growth Rate (Compared to Same Week in 2012) *NEW***

This weekly growth rate calculation is a useful indicator to look at weekly growth in registrations relative to average weekly registrations in 2012. This growth rate compares total weekly registrations (Sunday – Saturday) to weekly registrations for the corresponding week in 2012. The field is updated every Sunday morning at 12 am EST.
Equation:

Weekly growth rate (compared to same week in 2012) = \frac{(X-Y)}{Y} \times 100

Where X = Total weekly registrations
Y = Total weekly registrations for same epi week in 2012

**Average satisfaction rating for all users who responded to survey question (n=XXX)**

The average rating shown here is calculated as an average across all respondents since launch on 2/2/2010 who responded to the following survey question: “*How helpful was the text4baby service during your pregnancy? Reply with a number between 1 and 10, where 1 is the lowest rating and 10 the highest.*”

From launch – March 3rd, 2013, users in the pregnancy and new baby protocols were sent the survey question one month prior to due date and one month after birth date, respectively. Starting March 4th, only users in the pregnancy protocol will receive the survey question roughly 1 month prior to their due date. The number in the parentheses (n=) is the total number of respondents to this survey question since launch on 2/2/2010.

**Percent of users since launch (2/2/2010) who would refer a friend (n=XXX)**

The percentage shown here is calculated based on responses of users since launch who responded to the following survey question: “*Would you recommend the text4baby service to someone else who is pregnant or a new mother? Reply 1 for Yes and 2 for No.*”

From launch – March 3rd, 2013, users in the pregnancy and new baby protocols were sent the survey question 1 month prior to due date and one month after birth date, respectively. Starting March 4th, only users in the pregnancy protocol will receive the survey question roughly 1 month prior to their due date. The number in the parentheses (n=) is the total number of respondents to this survey question since launch on 2/2/2010.

**Percent of users who asked Dr./midwife about info in text4baby message (n=XXX) **

The percentage shown here is calculated based on responses of users who responded to the following survey question: “*Did you talk to your Dr. or midwife about any of the topics you read in text4baby messages? Reply 1 for Yes or 2 for No.*”

As of March 4th, 2013 users in the new mom protocol will receive this survey question in week 9 of the new baby protocol. The "n" value reflects the number of users who replied to the survey question since March 4th, 2013.
Percent of users who called resource from text4baby message \( (n=XXX) \) *NEW*

The percentage shown here is calculated based on responses of users who responded to the following survey question: “Did you call any of the phone numbers in your text4baby messages for more info or help? Reply 1 for Yes or 2 for No.”

As of March 4th, 2013 users in the new mom protocol will receive this survey question in week 9 of the new baby protocol. The "n" value reflects the number of users who replied to the survey question since March 4th, 2013.

Subscription Status of Users (Horizontal Bar Chart)
Displays the cumulative number of users since launch on 2/2/2010 through present who are and are not currently receiving messages

- Values can be:
  - ON: User is currently receiving protocol and/or ad-hoc broadcast messages.
  - OFF: User has texted in the word STOP and is not currently receiving any messages.
- Subscription status can be changed from OFF to ON or ON to OFF at any given point in time (i.e. if a user had another baby and decides to re-register, her status will be changed from OFF to ON).

Currently Subscribed Users by Protocol (Horizontal Bar Chart)
Displays the number of currently subscribed users by the type of messages they are receiving at the present time.

- Values can be:
  - Pregnancy = User is currently receiving pregnancy protocol and ad-hoc messages.
  - New baby \( \leq 1 \) = User is currently receiving new baby protocol and ad-hoc messages.
  - New baby \( >1 \) = User is currently only receiving ad-hoc messages

Currently Subscribed Users by Language (Horizontal Bar Chart)
The total number of currently subscribed users by language protocol they are in at the present time.

- Values can be:
  - English = User is currently subscribed to English protocol messages.
  - Spanish = User is currently subscribed to Spanish protocol messages.

Users since launch (2/2/2010) by trimester of pregnancy upon registration (Horizontal Bar Chart)
The total number of users, since launch, who registered directly into the pregnancy protocol the first time they registered for text4baby, split by the trimester in which they registered.
This field is calculated based on the due date the user entered during first registration in the pregnancy protocol. Users who first registered in the new baby protocol and later re-subscribed in the pregnancy protocol are not included in this chart.

• Trimester is defined as follows:
  o 1st: Week 1 of pregnancy through end of week 13 of pregnancy (Note: Users who register before week 7 of pregnancy will not receive any content messages until the 5th week, when the user receives 2 weeks of “Starter Pack” messages. There are no content messages before week 7 of pregnancy.)
  o 2nd: Beginning of week 14 of pregnancy through end of week 27 of pregnancy
  o 3rd: Beginning of week 28 of pregnancy through end of week 40

**T4b Dashboard – Outreach Partners (Tab 2)**

**DATASET 1 – Participant’s View for HMHB Outreach Partners**
The outreach partners dataset includes individual-level registration data. Some of the data elements, including protocol, subscription status, currently receiving protocol messages, zip code, county and state, represent MOST RECENT interaction with the system. Thus, if a user has updated multiple times or cancelled and re-registered, only the updated information will be showed for these data elements. ID, date of registration, channel of registration, language, type, and referral URL, do not change and will always reflect data collected at first registration.

The data elements listed below are visible in the sidebar tab labeled “Participants for Outreach Partners”.

**ID**
- ID is a unique number assigned to each user (unique cell phone number) upon first registration.

**Protocol**
- The message protocol that the user is enrolled in is based on either the due date or birth date provided during the most recent interaction with the system.
  o If the date provided is in the past (relative to date of registration or update), it is considered a birth date and user is enrolled in the new baby protocol.
  o If the date is in the future (relative to date of registration or update), it is considered a due date and user is enrolled in the pregnancy protocol.
  o The date provided cannot be more than 9 months in the future (for pregnancy) or more than 12 months in the past (for new baby), or an error message is sent during registration. A user cannot register without a valid date.
- Value can be “Pregnancy” or “New Baby.” This is a text value.
Note: If the user unsubscribed, the value displayed will reflect the protocol the user was in when they unsubscribed. This value is present for all users since program launch (2/2/2010).

Subscription Status
- Whether a user is or is not currently receiving SMS messages
- Values can be:
  - ON: User is currently receiving protocol and/or ad-hoc broadcast messages.
  - OFF: User has texted in the word STOP and is not currently receiving any messages.
- Subscription status can be changed from OFF to ON or ON to OFF at any given point in time (i.e. if a user had another baby and decides to re-register her status would change from OFF to ON).

Currently Receiving Protocol Messages*NEW*
- Whether or not a user is currently receiving protocol messages (pregnancy or new baby) sent three times a week.
  - Values can be:
    - YES: User is currently receiving either pregnancy or new baby protocol messages
    - NO: User has a baby >1 based on most recent date of birth and is only receiving ad-hoc broadcast messages
- Filter by currently receiving protocol messages = “yes” and protocol = “new baby” to calculate how many users in the new baby protocol are receiving protocol messages vs. how many users are not receiving protocol messages and only receiving adhoc messages.

Language
- The message protocol the user is enrolled in based on the keyword (BABY for English and BEBE for Spanish) submitted during first registration. Value is English or Spanish. This is a text value.
- All English protocol messages are provided in English. All Spanish protocol messages are provided in Spanish.
- This value is present for all users since program launch (2/2/2010).

Type
- The type of text4baby user reported by the user during first registration. Options include: (1) Pregnant, (2) New Mom, (3) Dad/Parent, (4) Relative/Friend, and (5) Healthcare Provider/Observer.
- This question was implemented starting with users who registered on 2/29/2012. In August 2012 we asked all currently subscribed users who registered prior to 2/29/2012 to provide their user type. If “type” is blank this is because a user was not
currently subscribed in August 2012 when the adhoc message was sent or because they did not respond to the adhoc question.

**Registration Date**
- The date that a user first enrolled in the text4baby program, via SMS, web or API.
- This is a date value in the format (MM, DD, YYYY).

**Channel of registration *NEW***
- Channel of registration captures how a user first registered for text4baby.
- Values include:
  - SMS: Users who registered via shortcode 511411
  - Web: Users who registered online
  - API: Users who registered with a third party who provided registration data to text4baby

**Zip Code**
- The five-digit zip code value that a user entered during the most recent interaction with the system.
- This data must be entered in five-digit format for a user to complete registration and enroll in text4baby.
- This is a numeric value. (Note: formatting errors in Excel may cause the number to appear as a four-digit value if it begins with or ends with “0”. Please format the cell to “Special – Zip Code”.)
- **Note:** For zip codes in which there are five or fewer enrollees, the zip code value will be hidden. This is to meet requirements around protecting the Personal Health Information of the users. County and State will still be shown.

**County**
- The geographic county in which the user is located, upon most recent registration in text4baby. This value is calculated based on the Y2IP code entered during the registration process, using the Y2IP code database from zipcodedownload.com.
- If one zip code represents more than one county, the system selects the first county on the list.

**State**
- The geographic state in which the user is located, upon most recent registration in text4baby. This value is calculated based on the ZIP code entered during the registration process, using the ZIP code database from zipcodedownload.com.

**Referral URL**
- Referral URL captures the URL used by users who register for text4baby via the web (first registration only). If no URL is displayed then the user registered for text4baby via SMS or API (not via the web).
• Values can be:
  o http://text4baby.org/ = A user who registered using the shadow box on the text4baby website.
  o http://www.text4baby.org/index.php/sign-up?url= [insert Outreach Partner website URL] = A user who registered using a registration button that has been placed on an Outreach Partner website.

**DATASET 2 – Participant’s View by Referring URL**
The Referral URL tab includes registration data for users who first registered for text4baby via the web; data for those who registered via SMS or API are NOT displayed. This tab is helpful if you are interested in looking at data specific to users, who register via the web, rather than downloading the large set of data available in the Participants for Outreach Partners tab and then filtering the data to focus on web registrations. All data elements included in this tab are defined in the section above (“Participants for Outreach Partners Data Definitions”) with the exception of channel of registration defined below:

**DATASET 3—Referral Survey *NEW***
The referral survey tab includes data for users who were asked a referral survey (first implemented with users who registered on 2/20/2012) to understand the referral sources of users who enroll in text4baby. This dataset includes records for all users who were asked the referral survey, including those who do not respond to the survey. All data elements included in this tab are defined in the section above (“Participants for Outreach Partners Data Definitions”) with the exception of referral source, submission status and submission date, which are defined below:
Referral Source

- Responses from users who registered from 2/20/2012 until 02/22/2013, will correspond with the following first referral survey: How did you hear about T4B? Reply 1: Friend/family; 2: Dr./nurse; 3: Health plan; 4: Health dept; 5: TV/radio/newspaper; 6: Online; or 7: Other.
- Responses from users who registered from 2/23/2013 - present, will correspond with the following second referral survey: How did you hear about T4B? Reply 1: Friend/family; 2: Dr./nurse; 3: Health plan; 4: Health dept; 5: TV/radio; 6: Magazine; 7: Facebook/Twitter; 8: Other
- Values include:
  - For users who registered from 2/20/2012–02/22/2013: Friend/family; Dr./nurse; Health plan; Health dept; TV/radio/newspaper; Online; Other.
  - For users who registered from 2/23/2013–present: Friend/family; Dr./nurse; Health plan; Health dept; TV/radio; Magazine; Facebook/Twitter; Other
- If you are interested in these data, filter by date of registration and analyze separately.

Submission Status

- Submission status captures whether or not a user responded to the survey question.
- Values include:
  - Draft: User has not responded to the survey question
  - Submitted: User has submitted a response to the survey question. All users with submission status = “submitted” will have response recorded for the referral source data element

Submission Date

- The date the referral survey was sent for users who did not respond to the survey (submission status = draft) or the date the referral survey was answered by users who provided a response (submission status = submitted).
- This is a date value in the format (MM, DD, YYYY).

Chart Definitions (Tab 3)

This tab includes 3 charts with national-level data. Please note that these charts are very slow to load. The definition of each chart found in the tab labeled “Charts” is provided below.

CHART 1 - Users Registered Per Hour for Past 30 Days

This chart shows the number of users who have registered within a given hour over the past 30-day period. Users are only counted once on the first date of registration. If a
user was previously registered in text4baby and subsequently re-subscribes, this user is not counted again. The X-axis shows the Date and Hour in EST. The Y-axis shows the Number of Users.

**CHART 2 - Users Who Have Registered Per Day, Since Text4baby Launch on 2/2/2010, including Cumulative Total Users Who Have Registered Since Launch**

This chart shows the number of users who have registered within a given day (defined as ending at 11:59 pm EST), since text4baby launched on 2/2/2010. Users are only counted once on first date of registration. If a user was previously registered in text4baby and subsequently re-subscribes, this User is not counted again. This chart also shows the cumulative number of users who have enrolled in text4baby over time, since the program launch on 2/2/2010. The X-axis shows the Date. The Y1-axis shows the Number of Users. The Y2-axis shows the Cumulative Total Users Since Launch on 2/2/2010.

**CHART 5 - Total Users Distributed by Week of Pregnancy Upon Original Registration in Pregnancy Protocol, Since Text4baby Launch on 2/2/2010**

This chart shows the total users distributed by week of pregnancy upon first registration into the Pregnancy protocol since text4baby was launched on 2/2/2010. The week of pregnancy is calculated based on the Due Date entered at the time of first registration. This chart does not take into account any updates to the due date that may have been provided by a user subsequent to first registration, nor any re-subscriptions into the Pregnancy protocol by a user. The X-axis shows the Week of Pregnancy from 1-40 weeks. The Y-axis shows the Number of Users.