Date: August 31, 2018

To: All Members

From: Board of Directors

Re: 2018 Outreach Evaluation for the Excavator Direct Mail Program

In the summer of 2018, an evaluation of the outreach of the Excavator Direct Mail Program was performed. The validation work was done by JKL Pipeline Consulting. A detailed report from JKL Pipeline Consulting dated August 15, 2018 is attached at the end of this letter.

The mailing in California primarily uses data from the CA State Licensing Board. In Arizona, a combination of InfoUSA, Acxiom and data from the AZ State License Board is used for the mailing. All other states use data from InfoUSA and Acxiom. Local data for municipalities from MCH is also included in all states. The data from MCH was not included in this evaluation because it was reviewed with the Public Officials program. The full listing of SIC codes for each area is included in the program outline.

Data from six list sources was evaluated: InfoUSA, BB Direct, National Data Group, Acxiom and State Business License data in California and Arizona. In depth evaluations were performed on two test counties: Sutter County, California and Mohave County, Arizona. The report indicates that the percent of stakeholders reached is approximately 67% in California, 89% in Arizona and approximately 35 - 45% elsewhere. These percentages are based on the total number of valid unique records determined from all sources compared to the sources used in the 2018 mailing (excluding MCH).

Based on the findings contained in the report, the Board is recommending the following actions be taken:

1. Eliminate California contractors with an "inactive" license status from the mailing. This is approximately 21% of the total California mailing.

2. Add the full list of SIC codes from the commercial data (InfoUSA and Acxiom) into the California mailing along with the "active" CSLB data. This will add back in approximately 37% (29% from InfoUSA and 8% from Acxiom), increasing the total quantity by approximately 16% and increasing the total outreach in California to approximately 87%.

3. Initiate a new research effort to identify other data sources to add to the ones being used now.

The recommendations are to be presented to the members at the Annual Meeting in September and on future monthly member calls. Confirmation from members regarding the recommendations and additional determinations will be documented in the minutes of those meetings.

<h1 style="text-align:center">MEMORANDUM</h1>

To: Jeff Farrells, Executive Director
     Pipeline Association for Public Awareness

From: Josie Long,
     JKL Pipeline Consulting, LLC.

Subject: List Verification Project

Date: August 25, 2018

Background

The Pipeline Association for Public Awareness (PAPA) and its member pipeline operators are required to follow American Petroleum Institute, Recommended Practice 1162 (API RP 1162). RP 1162 provides requirements for communicating with stakeholders about pipeline safety. To meet excavator requirements in RP 1162, PAPA annually communicates pipeline safety information with excavators through direct mail.

PAPA contracted with JKL Pipeline Consulting (JKL) to verify and validate six new and existing list sources (InfoUSA, Acxiom, BB Direct, National Data Group and State Contractor License data in California and Arizona) to determine the extent of coverage of each of the sources. The goal of the evaluation was to determine the percentage of excavator businesses that are valid records, calculate the number of businesses currently reached, and determine if additional list sources could improve future mailings.

For the purposes of this project, Sutter County, California and Mohave County, Arizona were selected for the analysis. JKL ordered lists from each of the list providers based on specific SIC codes and business license codes outlined in PAPA's *Excavator Program Overview* document. A separate analysis was conducted in each test county due to differences in how mailings lists are created in Arizona versus California. The California mailing includes the California Contractor License data while the Arizona mailing includes a combination of InfoUSA, Acxiom and Arizona Contractor License data.

METHODOLOGY

The analysis included multiple data validation techniques and a set of data validation rules to measure the accuracy, completeness and reach of each data source.

The following outlines the process used to validate records:

1. Six separate list sources were purchased containing excavators in Sutter County and Mohave County: InfoUSA, BB Direct, Acxiom, National Data Group and State Business License data in California and Arizona. Each excavator list was pulled from a set of SIC codes and business codes that PAPA uses to compile annual excavator mailings and contained at a minimum company name, mailing address, phone number and SIC code.

2. After each list was purchased, the data was integrated and normalized into a common database.   Records received a unique numerical identifier in order to maintain data integrity.

3. The following validation rules were established to determine if a record was valid:
   - If at least two of the lists contained the same entity based on company name and address, the record was considered valid.
   - If the California Contractor License data listed the entity as "Inactive," the record was considered out of scope for the analysis.
   - If the California Contractor License data listed the entity as "active," the record was considered a valid record.
   - If entities had certain identifiers in their name that were not likely to excavate (i.e., roofing companies, power washing companies, tiling companies, painting companies, etc.) they were considered "out of scope."
   - If records had the same company name, different addresses where one record had a physical address and the other had a PO Box, but phone numbers matched, the records were considered valid.
   - If records had one common field, but company name, address, or phone did not match, additional research was used to verify the record

4. A third party was used to validate the remaining records using a blend of phone calls and web research. Infogroup was selected to first conduct a phone interview with each unverifiable business. If the business was unable to be reached by phone, Infogroup used web research to verify the remaining records.
   - Infogroup called 1,375 business records to verify company name and address. In instances where the business was reached and the company name and address were confirmed, the record was considered valid. In instances where the company name and/or address was inaccurate or incomplete, Infogroup updated the record.
   - Remaining business records that were unable to be validated through phone calls were then attempted to be validated by Infogroup's research team. The team validated records by locating their business website address. If a URL was located, the record was considered validated and all contact information was updated if needed.

RESULTS

The following results were found through the validation analysis:

- Based on analysis, BB Direct was deleted as a source for the analysis. JKL found that 97% of the BB Direct list matched the InfoUSA list. When contacted, they noted that their list utilized InfoUSA data. Therefore, it was decided that the list source was not an appropriate way to validate businesses.

- JKL conducted a separate analysis in Sutter and Mohave counties. Several different analyses were performed in each county to understand list accuracy, list completeness, mailing reach, invalid rates and program cost impacts.

    o **List accuracy** measures how many valid excavators are in each data source compared to the total number of excavators included in the original list. List sources that have high list accuracy indicate that the list is of high quality.
    o **List completeness** measures the total number of valid excavators compared to the number of valid excavators on each list. Lists with high levels of completeness result in a more comprehensive dataset. Lists with low levels of completeness indicate that additional data sources may need to be considered in order to achieve higher reach.
    o **Reach** measures the number of validated records from the current mailing compared to the total universe of validated records. Reach scenarios demonstrate the percent increase in reach rates when additional list sources are added to the excavator mailing.
    o **Invalid rates** indicate the number of out of scope and unconfirmed excavators that are outside the Associations program. High invalid rates in list sources result in over mailing to businesses.
    o **Program cost impacts** take into account the quantity of both valid and invalid businesses since the current mailing does not have a mechanism to filter out invalid businesses. Cost scenarios calculate how additional data sources impact program costs.

Arizona Analysis

- **List Accuracy**: The number of validated records compared to the total list records, minus the out of scope records, represents the level of accuracy within the list. Based on the analysis by source, the Arizona Contractor License data has the highest accuracy. InfoUSA had an 78% accuracy rate, followed by Acxiom. National Data Group had the lowest percent accuracy, but a fairly high number of validated records. This in part is due to their large number of total records, which is more than any other source.

List Source	Total Records	Invalid Records	Validated Records	Accuracy Rate
Arizona Contractor License	751	5	746	99%
InfoUSA	640	143	497	78%
National Data Group	1,461	856	605	41%
Acxiom	421	95	326	77%

- **List Completeness**: To evaluate the completeness of each list, it was assumed that deduping the "Validated Records" reflects 100 percent of the total excavator universe. In total, there were 1,307 unique businesses that were validated through the analysis in Mohave, Arizona. Looking at the percentage of valid records in each list compared to the total valid universe, Arizona Contractor License data had the highest completeness at 57 percent, followed by the National Data Group at 46 percent. InfoUSA and Acxiom had the lowest completeness rates.

List Source	Total Validated	Validated Records	Completeness Rate
Arizona Business License	1,307	746	57%
InfoUSA	1,307	497	38%
National Data Group	1,307	605	46%
Acxiom Data	1,307	326	25%

- **Current Mailing Reach**: The Arizona mailing incorporates three commercially available lists. The base list is compiled using InfoUSA and Acxiom data and supplemented with Arizona Contractor License data. The Arizona Contractor License data is refreshed yearly and only includes active excavators. Reach rates were calculated by comparing the total number of validated excavators to the combined InfoUSA data, Acxiom and Arizona Contractor License list data. Based on the analysis, PAPA reached 89 percent of excavators in Mohave county.

InfoUSA Data	497
Add Unique Acxiom Data	86
Add Unique Arizona Contractor License List	580
Total	1,163
Total Validated Records	1,307
Percent Reach	**89%**

- **Invalid Records**: Each of the five sources contained a number of invalid records. Invalid records were categorized as either being out of scope for the analysis or unconfirmed. Records were deemed out of scope if their business name suggested their work was unrelated to excavation work, such as dry walling, roofing, painting, etc. Invalid records that were unable to be validated from multiple list sources, phone calls or web searches

were considered unconfirmed. When analyzing the total invalid records in Mohave, Arizona to the total list, approximately 45 percent of the unique records were invalid. When comparing total invalid records by list source, The National Data Group list contained the most invalid records of any list source.

List Source	Out of Scope	Unconfirmed	Invalid Rate
Arizona Business License	5	-	0.5%
InfoUSA	59	84	13%
National Data Group	343	513	78%
Acxiom Data	17	78	9%

- **Cost Analysis**: The current mailing includes Arizona Contractor License data, InfoUSA and Acxiom data. Based on the combination of commercial list sources and the number of records, the Association mails to approximately 1,406 businesses. This includes the total valid and invalid records. At a program cost of $1.10 per mailing, the total cost of the commercial lists excluding other add-ins is $1,546.60. If the Association were to add the National Data Group to the mailing, the mailing would increase to a quantity of 2,406 records and result in a 70 percent increase in program costs.

	Quantity	Program Cost
InfoUSA + Acxiom +AZ List	1,406	$1,546.60
InfoUSA + Acxiom + AZ List + NDG	2,406	$2,646.60

California Analysis

- **List Accuracy**: The number of validated records compared to the total list records, minus invalid records, represents the level of accuracy within the list. Based on the analysis by source, the Acxiom data, California Business License data and InfoUSA had similar accuracy rates. National Data Group had the lowest percentage of accuracy.

List Source	Total Records	Invalid Records	Validated Records	Accuracy Rate
California Business License	542	129	413	76%
InfoUSA	252	61	191	76%
National Data Group	645	359	286	44%
Acxiom Data	122	28	94	77%

- **List Completeness**: To evaluate the completeness of each list, it was assumed that deduping the "Validated Records" column reflects 100 percent of the total excavator universe. In total, there were 620 unique businesses that were validated through the

analysis in Sutter, California. Based on this analysis, the California Contractor License data had the highest completeness at 67 percent, followed by the National Data Group and InfoUSA. Acxiom had the lowest completeness rate.

List Source	Total Validated	Validated Records	Completeness Rate
California Business License	620	413	67%
InfoUSA	620	191	31%
National Data Group	620	286	46%
Acxiom Data	620	94	15%

- **Current Mailing Reach**: The California mailing only includes the California Contractor License list. Comparing the California Contractor License list to the total 620 valid records, results in a 67 percent reach rate. Adding additional list sources improves the total number of unique valid records and increases the mailing's reach. The Association can achieve an 87 percent reach by adding InfoUSA and Acxiom list sources or can achieve a 100 percent reach of validated records by adding the National Data Group list.

	Records	Reach
California Contractor License List	413	67%
First adding InfoUSA unique records	+107	84%
Second adding Acxiom unique records	+18	87%
Third adding NDG unique records	+82	100%

- **Invalid Records**: Each of the five sources contained a number of invalid records. Invalid records were categorized as being out of scope for the analysis if they were out of scope, unconfirmed or without a valid California license. Records were deemed out of scope if their business name suggested their work was unrelated to excavation work, such as drywall, roofing, painting, etc. Also, it should be noted that the California Contractor License data includes active and inactive excavators. Records that were unable to be validated from other lists, phone calls or web searches were considered unconfirmed. When comparing the invalids to the total unique records, approximately 47 percent of records were invalid. When comparing the list sources and invalid rates, the National Data Group list has the highest number of invalid records.

List Source	Out of Scope	Unconfirmed	Inactive License	Invalid Rate
California Business License	7	-	122	22%
InfoUSA	28	28	1	11%
National Data Group	163	192	4	62%
Acxiom Data	3	24	1	5%

- **Cost Analysis**: The current mailing includes approximately 420 businesses from the California Contractor License list. This includes the total number of valid and invalid records. At a program cost of $1.10 per business, the total cost of the commercial list excluding other add-ins is $462 in Sutter, California. If the Association were to add in other list sources, the mailing would increase as follows:

	Quantity	Program Cost
CA List + InfoUSA	588	$646.80
CA List + InfoUSA + Acxiom	634	$697.40
CA List + InfoUSA + Acxiom + NDG	1075	$1,182.50

Reach Rate Calculation: InfoUSA and Acxiom List Sources

The majority of county excavator lists are compiled using a combination of InfoUSA and Acxiom data. Based on the analysis, in areas where excavator lists are compiled using InfoUSA and Acxiom data sources, the reach is between 35 percent (Sutter County, CA) to 45 percent (Mohave County, AZ). This is calculated by dividing the combined unique counts of InfoUSA and Acxiom valid records to the total number of valid records in each county.

	InfoUSA and Acxiom Unique Records	Total Valid Records	Percent Reach
Sutter County, CA	215	620	35%
Mohave County, AZ	583	1,307	45%

Pages 8 and 9 show the calculations that were developed based on the validation analysis in each county.

Summary - Mohave County, Arizona

	Quantity	Percent	Cumulative%	Cost
Total Unique Records	**2370**	**100.0%**		
Valid Records	**1307**	**55.1%**		
InfoUSA	497	38.0%	38.0%	$546.70
Acxiom unique adds	86	6.6%	44.6%	$94.60
AZ Licensed unique adds	580	44.4%	**89.0%**	$638.00
NDG unique adds	144	11.0%	100.0%	$158.40
Invalid Records	**1063**	**44.9%**		
InfoUSA - Out of Scope	59	5.6%		$64.90
InfoUSA - Unconfirmed	84	7.9%	13%	$92.40
Acxiom Data - Out of Scope	17	1.6%		$18.70
Acxiom Data - Unconfirmed	78	7.3%	9%	$85.80
AZ Licensed List - Out of Scope	5	8.5%	0.5%	$5.50
NDG - Out of Scope	343	32.3%		$377.30
NDG - Unconfirmed	513	48.3%	78%	$564.30
*invalid records are not unique	1099		100%	

List Combinations

	Quantity	Cumulative%	Cost
2018 Mailing (for reference)	1638	89.0%	$1,801.80
InfoUSA + Acxiom Data + AZ Li	1406	89.0%	$1,546.60
InfoUSA + Acxiom Data + AZ Li	2406	100.0%	$2,646.60

Summary - Sutter County, California

	Quantity	Percent	Cumulative%	Cost
Total Unique Records	**1161**	**100.0%**		
Valid Records	**620**	**53.4%**		
CA list (active)	413	66.6%	**66.6%**	$454.30
InfoUSA unique adds	107	17.3%	83.9%	$117.70
Acxiom unique adds	18	2.9%	86.8%	$19.80
NDG unique adds	82	13.2%	100.0%	$90.20
Invalid Records	**541**	**46.6%**		
CA List - Inactive	122	22.6%		$134.20
CA List - Out of Scope	7	1.3%	22.4%	$7.70
InfoUSA - Out of Scope	32	5.9%		$35.20
InfoUSA - Unconfirmed	28	5.2%		$30.80
InfoUSA-License Inactive	1	0.2%	11%	$1.10
Acxiom - Out of Scope	3	0.6%		$3.30
Acxiom - Unconfirmed	24	4.4%		$26.40
Acxiom - License Inactive	1	0.2%	5%	$1.10
NDG - Out of Scope	163	30.1%		$179.30
NDG - Unconfirmed	192	35.5%		$211.20
NDG - License Inactive	4	1%	62%	$4.40
*invalid records are not unique	577		100.0%	

List Combinations

	Quantity	Cumulative%	Cost
2018 Mailing	570	66.8%	$627.00
CA List	420	66.8%	$462.00
CA List + InfoUSA	588	83.8%	$646.80
CA List + InfoUSA + Acxiom Data	634	86.5%	$697.40
CA List + InfoUSA + Acxiom Data +NDG	1075	100.0%	$1,182.50