Conventions

Below are conventions that we would expect during typical data delivery.

File Delivery - Typical buyer partners expect S3 or Google Cloud file delivery in a flat file format. Multiple Files should be sent breaking up each file to be under 1M rows in the below examples:

- Multiple files namespaced into a folder within an S3 bucket by date, following the naming convention: <D>_<M>_<Y>/<X>-<Batch Number>-<Part Number>.csv.gz. Multiple single files will be sent throughout the day with no size limits.

  <D>, <M>, and <Y> should be the day, month, and year the file is delivered, respectively. <X> can be any file name identifier. <Part Number> is the file breakup id for a batch. <Batch Number> is the file id for a single delivery. Example: *30_05_2020/pool-foundation-outgoing-1-00001.csv.gz*

  File Full Address Example: *s3://poolfoundation-incoming/30_05_2020/poolfoundation-mobile-outgoing-1-00001.csv.gz*

Compression - Files should have gzip compression when sent.

Format - Files will be formatted with Comma Separated Values specification, using a backslash escape character when necessary.

Headers - Files should not include header information, but should match the data schema provided by your Data Union.

S3 Permissions - S3/Google Cloud credentials should be provided to you with S3:Get, S3:List, S3:PutObject.

Frequency

Please communicate delivery times with the Pool Foundation to help understand and debug data. We will generate automatic alerts if data is not delivered within the specified time to help diagnose or debug issues.

Fill Rates

Please communicate fill rates for each data point that you typically expect. For example, hashed email data should have a 100% Fill Rate on the SHA or MD5 values, etc., but may have a low fill rate

(1-10%) on the Latitude, Longitude, etc. The Agreement will define the data points that are required and the data points that will be passed when available for each data type.

Data Specification Schema

The Data Specification schema should be a spreadsheet that describes the datapoints being delivered for each row. An example of a Schema is below. If there are any issues or discrepancies, please let us know.

Data Point	Definition	Type
Mobile ID	This can be Android Ad ID or Apple's IDFA. This must be delivered in clear-text (unhashed). An IDFA Mobile ID hash is represented by a 32-character hexadecimal string where the alphabetic characters are uppercase in the hashes, broken up by 4 dashes and looks like 8-4-4-4-12. An AAID Mobile ID hash is represented by a 32-character hexadecimal string where the alphabetic characters are lowercase in the hashes, broken up by 4 dashes and looks like 8-4-4-4-12.	String
Device Type	The type of device that the MAID was collected from. "IDFA" for ios and "AAID" for Android. Answers should be all uppercase alphabetic letters.	String
IP Address	This value is the public /external IP address of the device that made the request, received on the server side. It should be in IPv4 format but we do accept IPv6 format, e.g. IPv4 - "66.249.66.53. IPv6-2001:db8:3333:4444:5555:6666:7777:8888	String
Date Timestamp	Timestamp of the moment the event was recorded. It is expressed in Unix UTC seconds. For ex, 2/15/2018 23:33:31 = 1518737611	Integer
Latitude	With at least four (4) decimal points and max eight (8) decimal points of precision. Precision should reconcile with accuracy for each individual location point. Integers only. Example: "32.7787"	Double
Longitude	With at least four (4) decimal points and max eight (8) decimal points of precision. Precision should reconcile with accuracy for each individual location point. Integers only. Example: "-96.8217"	Double
Horizontal Accuracy	This is the horizontal accuracy of the location point in meters. Horizontal accuracy greater than 50 meters is considered too vague to use for location targeting; we prefer to have accuracies less than 50 meters for users. Use integer values rounded to the nearest integer. Example: "6"	Double

Vertical Accuracy	This is the vertical accuracy of the location point in meters. If the vertical accuracy is not available, this field should be left blank. Use integer values rounded to the nearest integer. Example: "2"	Double
Altitude	Altitude should return an integer value that represents the elevation above sea level in meters. Integers only. Example: "35"	Double
Accelerometer	Moving average of the device over 5 seconds. The format should be x: double, y: double, z: double. Integers only and separate each section with "-". Example: "3.00-4.50-9.00"	Double
Dwell Time	Provided in milliseconds. How long a device remains in one place before moving. Integers only. Example: "224000"	Time
Wireless Carrier	This should be the name of the wireless carrier. All alphabetic letters should be uppercase and spaces will be separated by "-". Examples: "T-MOBILE" "AT&T" "VERIZON"	String
Wireless Roaming	This should be the name of the wireless roaming carrier if the user is roaming. All alphabetic letters should be uppercase and spaces will be separated by "-". Examples: "T-MOBILE" "AT&T" "VERIZON"	String
Wireless Network	This should be the name of the wireless network, for example: "ACTV". If the network is not available this value should be left blank. All alphabetic letters should be uppercase and spaces will be separated by "-"	String
Sim MCC	Sim Mobile Country Codes (MCC) are used in wireless telephone networks (GSM, CDMA, UMTS, etc.) in order to identify the country which a mobile subscriber belongs to. This is integers only. Example: "311" "310" "204"	String
Sim MNC	Sim Mobile Network Codes (MCC) are used in wireless telephone networks (GSM, CDMA, UMTS, etc.) in order to identify the country which a mobile subscriber belongs to. This is integers only. Example: "480" "4" "210"	String
Bluetooth Device Name	Bluetooth devices connected or attached to the device (mobile-ID). All answers should have the alphabetic letters uppercase and integers. Any spaces should be separated by "-". Examples: "TOYOTA-COROLLA" "GOPRO-3278" "JBL-FLIP-2"	String
Wifi SSID	This is the network the device is connected to on wifi. All answers will include integers and alphabetic letters. Spaces are to be separated by "-". Example: "NETGEAR99" "xfinitywifi" "FIOS-8SXKP"	String
Wifi BSSID	Basic service set identifiers (BSSID) are used to describe sections of a wireless local area network or WLAN. It recognizes the access point or router because it has a unique address which creates the wireless	String

	network. Thus, BSSID is simply the MAC address of a wireless access point or also known as WAP. All answers should include lowercase alphabetic letters and integers. Examples: "7e:b0:66:02:22:1f" "94:10:3e:e4:91:40" "fc:5b:39:53:f7:cf"	
LC Raw Email	Raw lowercase consumer email address. Example: jondoe@gmail.com	String
LC MD5 Email	Any alphabetic characters in the username identifier must be converted to lowercase prior to hashing. When hashing, all identifiers must use the ASCII (single-byte) character set. Each identifier must be hashed only once using the MD5 hash algorithm, with no "salt" or "secret" value appended. An MD5 hash is a 16-byte value, which is represented by a 32-character hexadecimal string where all alphabetic letters are lowercase.	String
LC SHA1 Email	Any alphabetic characters in the username identifier must be converted to lowercase prior to hashing. When hashing, all identifiers must use the ASCII (single-byte) character set. Each identifier must be hashed only once using the SHA-1 hash algorithm, with no "salt" or "secret" value appended. A SHA-1 hash is a 20-byte value, which is represented by a 40-character hexadecimal string where all alphabetic letters are lowercase.	String
LC SHA256 Email	Any alphabetic characters in the username identifier must be converted to lowercase prior to hashing. When hashing, all identifiers must use the ASCII (single-byte) character set. Each identifier must be hashed only once using the SHA-256 hash algorithm, with no "salt" or "secret" value appended. A SHA-256 hash is a 32-byte value, which is represented by a 64-character hexadecimal string where all alphabetic letters are lowercase.	String
App Package Name	The app package name that has been extracted from the device, in its original format (e.g., "com.instagram.android"). Answers will include integers and alphabetic letters. This will be the app that the data is gathered from to populate this record of data.	String
Country Code	A two-character country code, using the standard ISO-3166-1 format. All alphabetic letters are to be uppercase only. Example, "US".	String