

Data Dictionaries

A comprehensive review of Batavia Research's company-centric data, data structure and derivative insights

Confidentiality Notice

This document is intended to be viewed only by recipients employed or affiliated with an organization with which an existing mutual non-disclosure agreement (NDA) is in place and active with Batavia Research. The following information is privileged under the terms of the executed NDA.

Introduction

Batavia Research (BR) serves business information and financial service providers with company-centric data that reduces the cost of data verification, improves accuracy, expands "thin-files", and augments risk and performance insights. This is achieved through in-house technology that discovers and gathers data from online sources to characterize a company's digital and physical presence and activities.

The following data dictionaries contain the company-centric attributes that BR collects and calculates. These attributes are collectively referred to as the "flat file". The flat file is derived from normalized and mapped raw data that is not reflected in these data dictionaries. The raw data is available upon request via supplemental dictionaries. Supplemental files include unmapped data native to each channel and unstructured content such as posts, media, or comments. Not all raw data is available due to the terms of service for a particular data source.

Flat File Categories

The flat file is subdivided into the following categories.

- Flat File Reference Information
- Flat File Analysis Flags
- Prime Company Insights
- Basic Company Information
- Organizational Data
- Operational Data
- Customer Review Data
- Website Information
- Social Media Channels
- Social Media Insights
- Content Insights
- Tools and Platforms

Each of these categories is further characterized in the tables below.

Data Dictionaries

The following data dictionary contains attributes that Batavia Research returns for each company record. The attribute labels are in the format native to BR's API and export templates. This data is mapped and normalized from raw, supplemental data not reflected in the following dictionary.

NOTE: Attributes with the prefix "dp" are values from a client's original records. This data is kept in the flat file to contrast the data we originate against a client's original data. It is also used for discrepancy analysis and attribute verification. These attributes are highlighted in green.

Object	Attribute Name	Description	Type
Flat File Reference Information			
company	dp_id	record key native to our client	num
company	br_company_id	internal ID used by BR	num
company	data_range	a minimum of 90 days is available	num
company	last_data_update	once a company is registered in BR, it is updated every 7 days	date time
Flat File Analysis Flags			
company	br_record_touched_flag	a record is touched if BR adds any data to it	txt
company	br_verification_flag	if any of a client's original attributes are verified AND there are no discrepancies	enum(verified, unverified, discrepancy)
company	br_channels_found_flag	one or channels associated with the company were detected	txt
Prime Company Insights			
company	br_company_operating_status	Prime insight based on the BRI logic	enum(open, closed, indeterminate)
company	br_cmi	"Content Marketing Index" score (0-100)	num
company	br_cmi_change	% change of cmi over time	num
company	br_last_activity_date	date of last post	date time
company	br_last_activity_channel	channel associated with the last post	Link
company	br_bri	Business Risk Indicator infers the going concern of a company (1-9)	num
company	br_bri_change	% change of bri over time	num
company	br_downgrade_alert_flag	partner can set alert based on BRI	num
Basic Company Information			
company	dp_company_name	name of company provided by partner's original record	string
company	br_company_name	company name discovered by BR	string
company	company_name_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	dp_street	street address provided by partner's original record	string
company	br_street	street address discovered by BR	string
company	street_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	dp_city	city provided by partner's record	string

Object	Attribute Name	Description	Type
company	br_city	city discovered by BR	string
company	city_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	dp_state	state(geo provided by partner's original record)	string
company	br_state	state (geo) discovered by BR	string
company	state_discrepancy	flags disposition of partner's records and BR's	enum(na, match, found, discrepancy)
company	dp_zip	zip provided by partner's original record	num
company	br_zip	zip discovered by BR	num
company	zip_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	br_lat	company's geo latitude	num
company	br_long	company's geo longitude	num
company	br_headquarters_address	company's headquarter address	string
company	dp_phone	phone provided by partner's original record	txt
company	br_phone	phone number discovered by BR	txt
company	phone_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	dp_fax	fax provided by partner's original record	txt
company	br_fax	fax number discovered by BR	txt
company	fax_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
company	dp_company_email	company's email provided by partner's original record	txt
company	br_company_email	company's email discovered by BR	txt
company	company_email_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
Organization Data			
company	br_parent_company	name of company's parent company	string
company	br_founded_date	founding of company	date
company	br_year_closed	if no longer operational, year of closure	date
company	dp_sic_code	sic code provided by partner	num
company	br_industry	industry label from BR	string
company	br_description	company description from BR	string
company	br_company_type	company's self-described label	string
company	br_employee_count	number range of employees	string
company	br_employee_count_change	change in number of employees	num
Operational Data			
company	br_hours_of_operation	company's operating house	string
company	br_payment_methods	accepted payment methods	string

Object	Attribute Name	Description	Type
company	br_price_range	relative price of services base on normalized native score	string
company	br_parking	indicates parking/car access at physical location	string
Customer Review Data			
company	br_review_count	number of consumer review for company of location	num
company	br_last_review_date	last date a review was left for a company or location	date time
company	br_rating	normalized rating	num
company	br_rating_change	% change in normalized rating	num
Website Information			
company	dp_website	company's website provided by partner	link
company	br_dp_website_state	state of company's website	enum(valid, invalid, na)
company	br_website	website discovered by BR	link
company	br_website_state	state of company's website (BR)	enum(valid, invalid, na)
company	website_discrepancy	disposition of partner's records against BR's	enum(na, match, found, discrepancy)
Social Media Channels			
company	primary_facebook_url		link
company	primary_facebook_unclaimed	indicates company has yet to "claim" their Facebook profile	link
company	primary_google_plus_url		link
company	primary_google_plus_unclaimed	indicates company has yet to "claim" their Google+ profile	link
company	primary_twitter_url		link
company	primary_youtube_url		link
company	primary_flickr_url		link
company	primary_foursquare_url		link
company	primary_instagram_url		link
company	primary_linkedin_url		link
company	primary_pinterest_url		link
company	primary_slideshare_url		link
company	primary_tumblr_url		link
company	primary_vimeo_url		link
company	primary_yelp_url		link
Social Media Insights			
company	br_total_channels	total number of channels where a company maintains an active presence	num
company	br_channel_change	% change in total number of channels	num
company	br_total_channel_followers	cumulative number of channel followers across all of a company's channels	num
company	br_followers_change	% change in follower growth	num

Object	Attribute Name	Description	Type
company	br_total_channel_posts	total number of posts (for life of channel)	num
company	br_posts_change	% change posting rate	num
company	br_total_channel_engagements	total number of engagements at the channel level (e.g. channels likes)	num
company	br_total_posts_engagement	total number of engagements at the posts level (e.g. comments, posts likes)	num
company	br_posts_engagements_change	% change in engagement	num
company	br_total_posts_comments	total post comments which are a subset of post engagement	num
company	br_post_comments_change	% change in post comments	num
company	br_dormant_channels	number of channels that are inactive in the last 90 days	num
company	br_dormant_channel_change	% change in dormant channels during the last 90 days	num
company	br_average_posts_per_channel	average posts per channels per day during the last 90 days	num
company	br_average_posts_change	% change in average posts per day	num
company	br_post_frequency	number of posts per days from a company or brand	num
company	br_post_frequency_change	% change in number of posts per day from a company or brand	num
company	br_media_variation	number of different media types used by a company (e.g. text, notes, video, picture, audio...)	num
company	br_engagements_per_posts	engagement ratio for posts	num
company	br_engagements_posts_change	% change in engagement ratio for posts	num
company	br_authors_count	number of users posting content for a company	num
company	br_authors_count_change	% change is authors	num
Content Insights			
company	br_most_popular_post	company's most engaged content	link
company	br_languages_used	types of languages associated with company's content and comments	string
company	br_social_handles	list of company's social handles (profile usernames)	string
company	br_hashtags	list of hashtags used in posts	string
company	br_mentions	list of social handles within posts and engagements	string
company	br_media_types	8 possible media types	enum(video, txt, photo, note, message, ...)
Tools and Platforms			
company	br_posting_tools	name of tool(s) used to post content to channels	string
company	br_analytics_tools	name of tool(s) used to perform website analytics	string
company	br_cms_platforms	name of platform(s) used to support website or blog	string

Page Left Blank

Batavia Research
1592 North Batavia
Orange, CA 92867
www.bataviaresearch.com
1.714.834.1200

David van der Roest
Partner Development & Cofounder
david@bataviaresearch.com
1.714.549.2446