FOOTFAL AND BRAND AFFINITY

- Points

Product Guide

Version: 2021

Audience Profiles - Footfall and Brand Affinity Points Product Guide

Version: 2021
Last Update: 18 October 2021

Table of Contents

1 – Getting Started

In this section

Audience Profiles - Footfall and Brand Affinity Points Overview

Locomizer's Audience Profiles - Footfall and Brand Affinity Products provide dynamic analysis of the spatial distribution of the general population (Footfall) and audience composition by interest (Brand Affinity) at any location at any time. This data is used by retailers, city planners, local authorities, financial services, and advertisers to discover audiences and their behavioural traits for critical business intelligence applications.

The Locomizer's Audience Profiles - Footfall and Brand Affinity Products consist of two datasets built using human mobility data collected from mobile devices where users have provided explicit consent for data to be used in this way. The product contains aggregated (Footfall) and abstract (Brand Affinities) data derivatives and adheres to local privacy regulations such as GDPR. The point data for location selection and user interest (affinity) profiling is supplied by Precisely, inc as a Points of Interest (POI) dataset extracted from World Points of Interest – Premium product. The POI dataset is updated monthly.

Footfall Dataset
This dataset provides the key footfall metrics at a point level (latitude and longitude of POI + 69-meter radius) where people spend time and/or money. The metrics are built using the mobile trace data generated by a representative panel of anonymous users. This includes the observed and extrapolated number of users who spent time at the point during a specific period, the number of their signals (observations) at the point, and the footfall data derivatives like reach, dwell-time, score. These metrics are provided for different movement (pedestrian and non-pedestrian) and visitation (residents, workers, transient) modalities. This dataset covers all needs for anyone to measure the popularity of a specific area and to perform comprehensive data analysis, build attribution and predictive models, and developed applications for retail analytics, city planning, marketing/advertising, financial forecasting and risk assessment etc.

Brand Affinity Dataset
This dataset complements the Footfall dataset and provides the quantitative measure of the real-world interest (affinity) of the users observed at a point level (latitude and longitude of POI + 69-meter radius). This includes hundreds of potential real-world interests including (but not limited to) shopping, leisure, sport, transport, entertainment, education. eating/drinking etc grouped at the brand level like KFC, Superdrug, Vue cinema, Virgin Active Gyms etc. The user interest (affinity) profiling methodology uses the supervised machine learning algorithm described in the Locomizer's patent "Interest Profile of a User of a Mobile Application".

Footfall and Brand Affinity Parameters Definitions

Number of Users
A number of unique users from a mobile data sample observed within a specified geographical area[1] during a specified period[2]

Number of Signals
A number of mobile phone signals generated by users within a specified geographical area[1] during a specified period[2]

Reach
An estimate of a percentage of unique users from a mobile data sample observed within a specified geographical area[1] during a specified period[2] in respect to a total number of unique users from a mobile data sample observed within a wider geographic area[3]

Dwell Time
An estimate of a percentage of mobile phone signals generated by users located within a specified geographical area[1] during a specified period[2] as a percentage of all mobile phone signals of these users generated within a wider geographical area[3] throughout the day (00-24)

Average Number of Signals
An average number of mobile phone signals generated by a user in a specified geographical area[1] during a specified period[2]

Proportion of Signals
A percentage of all mobile phone signals generated by users in a specified geographical area[1] during a specified period[2] to all mobile phone signals of these users generated within a wider geographical area[3] throughout the day

Footfall score
A function of the Reach and the Dwell Time used to rank a specified geographical area[1] during a specified period[2] with respect to the likelihood of finding a general user in a specified geographical area[1]

Footfall Score Normalized
Brand Affinity Score normalized on a scale of 0-100 %

Footfall Score Rank
Brand Affinity Score expressed as the percentrank on a scale of 0-1

Footfall Score SD
Brand Affinity Score expressed as the difference between the Footfall Score and the mean value of all Brand Affinity Scores in the daily dataset using standard deviation units

Extrapolated number of users
An estimate of the actual number of users in a specified geographical area[1] during a specified period[2].

Extrapolated number of signals
An estimate of the actual number of mobile phone signals generated in a specified geographical area[1] during a specified period[2]

Brand Affinity Score
A measure of the average Brand Affinity Score (real-world interest to physical locations, e.g., POIs) of users in a specified geographical area[1] during a specified period[2] using interest profiling algorithms described in the Locomizer's patent "Interest Profile of a User of a Mobile Application".

Brand Affinity Score Normalized
Brand Affinity Score normalized on a scale of 0-100 %

Brand Affinity Score Rank
Brand Affinity Score expressed as the percentrank on a scale of 0-1

Brand Affinity Score SD
Brand Affinity Score expressed as the difference between the Brand Affinity Score and the mean value of all Brand Affinity Scores in the daily dataset using standard deviation units

Brand Affinity Reach
An estimate of a percentage of unique Target users[4] from a mobile data sample observed within a specified geographical area[1] during a specified period[2] in respect to a total number of Target users[4] from a mobile data sample observed within a wider geographic area[3]

Brand Affinity Dwell Time
An estimate of a percentage of mobile phone signals generated by Target users[4] located within a specified geographical area[1] during a specified period[2] as a percentage of all mobile phone signals of these users generated within a wider geographical area[3] throughout the day (00-24)

The data in this product covers 24 hours a day and is categorized by calendar day.

Notes:

[1] POI with a specified radius

[2] Daypart

[3] Country

[4] Users from a mobile data sample with a positive Brand Affinity Score, which is measured during the previous month

Movement Modality

- Pedestrians: 0 – 5.0 km/h
- Non_pedestrians: 5.1 – 200 km/h

Visitation Modality

Residential classification is defined as identifying an area (using device level data) where a user has the highest dwell-time during the night hours (00.00-06.00 with a heavier weighting for 02.00-04.00) – this is used to infer the "common evening location (CEL)" or "home". The residential users are aggregated and output at the point level (POI + 69-meter radius).

Workers classification is defined as identifying an area (using device level data) where a user has the highest dwell-time during the working hours – this is used to infer the "common daytime location (CDL)" or "work". The worker users are aggregated and output at the point level (POI + 69-meter radius).

Transient classification is defined as identifying users who do not live and do not work in the area, consisting of local visitors and international visitors. International visitors are identified using tourist/country visitor classification algorithms, which segments the users based on the time they are seen in the country. The Transient users are aggregated and output at the point level (POI + 69-meter radius).

Dayparts

- Allday: 00.00-23.59
- Morning: 06.00 – 11.59
- Day: 12.00 – 17.59
- Evening: 18.00 – 23.59
- Night: 00.00 – 05.59

Footfall Dataset

	MOVEMENT MODALITY		
VISITATION MODALITY	All	Pedestrians	Non_Pedestrians
All[1]	Yes	Yes	Yes
Residents[2]	Yes	No	No
Workers[3]	Yes	No	No
Transient[4]	Yes	No	No

Brand Affinity Dataset

	MOVEMENT MODALITY		
VISITATION MODALITY	All	Pedestrians	Non_Pedestrians
All[1]	Yes	Yes	Yes
Residents[2]	No	No	No
Workers[3]	No	No	No
Transient[4]	No	No	No

Notes:

[1] **All** = Residents + Workers + Transient

[2] Residents of the point (POI + 69-meter radius) – a subset of residents of the UK

[3] Workers in the POI (POI + 69-meter radius) – a subset of workers of the UK

[4] Non-residents and non-workers of the point (POI + 69-meter radius)

Product Specifications

Geographic Coverage	United Kingdom
Geographic Detail	Radius of 69 meters
File Formats	Tab separated values (TSV)
Update Frequency	Weekly / Monthly according to contract

Field Names in Schema Tables

Field names in schema tables are documented in upper-case letters. The appearance of filed names in the actual product may differ from this convention.

2 - Data Schema

In this section

Audience Profiles – Footfall and Brand Affinity Points Data Schema

Data schemas may vary from country to country. Layout varies by country and by product.

Footfall Data Schema Table

Seq.	Field Name	Data Type (Length)	Description
1	CODE	STRING	Unique POI ID
2	RADIUS	INTEGER	Size of radius in meters
3	LATITUDE	FLOAT	Y value for POI centroid
4	LONGITUDE	FLOAT	X value for POI centroid
5	NAME	STRING	Primary/Registered name of the business
6	BRANDNAME	STRING	Precisely standardized brand name used by the business. BRANDNAME is only populated for certain countries and brands. Please refer to the Product Metrics for coverage information.
7	FRANCHISE_NAME	STRING	Franchise name
8	ISO3	STRING	Three-character ISO code of the country
9	AREANAME3	STRING	Name of the city where the business is physically located
10	AREANAME2	STRING	District name or equivalent. Each country has its own Administrative structure. AREANAME2 may not have an equivalent in every country
11	STREETNAME	STRING	Parsed street name
12	POSTCODE	STRING	Postal code where the business is physically located
13	FORMATTED_ADDRESS	STRING	A formatted address is an input address in a uniformly standard format as described by a set of attributes including

			Housenumber, Streetname, Streetname2, Areaname3 and Postcode
14	MAIN_ADDRESS_LINE	STRING	Address in a uniformly standard format including Housenumber, Streetname, Streetname2,
15	ADDRESS_LAST_LINE	STRING	Address in a uniformly standard format including Areaname3 and Postcode
16	HTTP	STRING	URL (Uniform Resource Locator) address of the business. URL availability may depend on nature of the business
17	BUSINESS_LINE	STRING	Description of the operations or activities of the business, which relates to the primary four-digit 1987 US SIC
18	SIC8_DESCRIPTION	STRING	8 Digit SIC description identifying a line of operations for a business at the most specific level.
19	MICODE	STRING	Precisely POI classification reserved set of MiCodes[1] which was agreed to be the 1099**** Code space as the "Reserved Space"
20	TRADE_DIVISION	STRING	Level1 POI category
21	GROUP_NAME	STRING	Level2 POI category
22	MAIN_CLASS	STRING	Level3 POI category
23	SUB_CLASS	STRING	Level4 POI category
24	GLOBAL_ULTIMATE_BUSINESS_NAME	STRING	Name of the ultimate company
25	HOURS	STRING	Daypart data was collected
26	DAY	DATE	Date data was collected formatted as ISO 8601:2004 YYYY-MM-DD
27	MONTH	INTEGER	Month data was collected
28	YEAR	INTEGER	Year data was collected
29	MOVEMENT_MODALITY	STRING	Classification of human movement

30	VISITATION_MODALITY	STRING	Classification of visitation
31	NUMBER_OF_USERS	INTEGER	Number of users
32	NUMBER_OF_SIGNALS	INTEGER	Total number of signals
33	DWELL_TIME	FLOAT	Dwell time
34	AVERAGE_NUMBER_OF_SIGNALS	FLOAT	Average number of signals per user
35	REACH	FLOAT	Reach
36	PROPORTION_OF_SIGNALS	FLOAT	Proportion of audience signals
37	FOOTFALL_SCORE	FLOAT	Footfall Score
38	FOOTFALL_SCORE_NORMALIZED	FLOAT	Normalized Footfall Score
39	FOOTFALL_SCORE_RANK	FLOAT	Footfall Score Percentrank
40	FOOTFALL_SCORE_SD	FLOAT	Footfall Score on a standard deviation scale
41	EXTRAPOLATED_NUMBER_OF_USERS	FLOAT	NUMBER_OF_USERS x (Mobile Device Holding Population / [Number of residents in the UK)
42	EXTRAPOLATED_NUMBER_OF_SIGNALS	FLOAT	EXTRAPOLATED_NUMBER_OF_USERS x AVERAGE_NUMBER_OF_SIGNALS_PER_USER

Brand Affinity Data Schema Table

Seq.	Field Name	Data Type (Length)	Description
1	CODE	STRING	Unique POI ID
2	RADIUS	INTEGER	Size of radius in meters
3	LATITUDE	FLOAT	Y value for POI centroid
4	LONGITUDE	FLOAT	X value for POI centroid
5	NAME	STRING	Primary/Registered name of the business
6	BRANDNAME	STRING	Precisely standardized brand name used by the business. BRANDNAME is only populated for certain countries and brands. Please refer to the Product Metrics for coverage information.
7	FRANCHISE_NAME	STRING	Franchise name
8	ISO3	STRING	Three-character ISO code of the country
9	AREANAME3	STRING	Name of the city where the business is physically located
10	AREANAME2	STRING	District name or equivalent. Each country has its own Administrative structure. AREANAME2 may not have an equivalent in every country
11	STREETNAME	STRING	Parsed street name
12	POSTCODE	STRING	Postal code where the business is physically located
13	FORMATTED_ADDRESS	STRING	A formatted address is an input address in a uniformly

			standard format as described by a set of attributes including Housenumber, Streetname, Streetname2, Areaname3 and Postcode
14	MAIN_ADDRESS_LINE	STRING	Address in a uniformly standard format including Housenumber, Streetname, Streetname2,
15	ADDRESS_LAST_LINE	STRING	Address in a uniformly standard format including Areaname3 and Postcode
16	HTTP	STRING	URL (Uniform Resource Locator) address of the business. URL availability may depend on nature of the business
17	BUSINESS_LINE	STRING	Description of the operations or activities of the business, which relates to the primary four-digit 1987 US SIC
18	SIC8_DESCRIPTION	STRING	8 Digit SIC description identifying a line of operations for a business at the most specific level.
19	MICODE	STRING	Precisely POI classification reserved set of MiCodes[1] which was agreed to be the 1099**** Code space as the "Reserved Space"
20	TRADE_DIVISION	STRING	Level1 POI category
21	GROUP_NAME	STRING	Level2 POI category
22	MAIN_CLASS	STRING	Level3 POI category

23	SUB_CLASS	STRING	Level4 POI category
24	GLOBAL_ULTIMATE_BUSINESS_NAME	STRING	Name of the ultimate company
25	HOURS	STRING	Daypart data was collected
26	DAY	DATE	Date data was collected formatted as ISO 8601:2004 YYYY-MM-DD
27	MONTH	INTEGER	Month data was collected
28	YEAR	INTEGER	Year data was collected
29	MOVEMENT_MODALITY	STRING	Classification of human movement
30	VISITATION_MODALITY	STRING	Classification of visitation
31	BRAND_AFFINITY_CATEGORY_NAME	STRING	Name of POI category for audience interest measurement (e.g., Precisely standardized brand name used by the business)
32	BRAND_AFFINITY_CATEGORY_MICODES	STRING	Precisely POI classification reserved set of MiCodes[1] which was agreed to be the 1099**** Code space as the "Reserved Space"
33	BRAND_AFFINITY_SCORE	FLOAT	Audience interest
34	BRAND AFFINITY_SCORE_NORMALIZED	FLOAT	Normalized Audience interest
35	BRAND AFFINITY_SCORE_RANK	FLOAT	Audience interest percentrank
36	BRAND AFFINITY_SCORE_SD	FLOAT	Audience interest on the standard deviation scale
37	BRAND AFFINITY_REACH	FLOAT	Target Reach
38	BRAND AFFINITY_DWELL_TIME	FLOAT	Target Dwell-time

Footfall Baseline Data Schema Table

Seq.	Field Name	Data Type (Length)	Description
1	DAY	DATE	Date data was collected formatted as ISO 8601:2004 YYYY-MM-DD
2	MONTH	INTEGER	Month data was collected
3	YEAR	INTEGER	Year data was collected
4	MOVEMENT_MODALITY	STRING	Classification of human movement
5	VISITATION_MODALITY	STRING	Classification of visitation
6	NUMBER_OF_USERS	INTEGER	Number of users
7	NUMBER_OF_SIGNALS	INTEGER	Total number of signals
8	AVERAGE_NUMBER_OF_SIGNALS	FLOAT	Average number of signals per user

Brand Affinity Baseline Data Schema Table

Seq.	Field Name	Data Type (Length)	Description
1	DAY	DATE	Date data was collected formatted as ISO 8601:2004 YYYY-MM-DD
2	MONTH	INTEGER	Month data was collected
3	YEAR	INTEGER	Year data was collected
4	MOVEMENT_MODALITY	STRING	Classification of human movement
5	VISITATION_MODALITY	STRING	Classification of visitation
6	BRAND_AFFINITY_CATEGORY_NAME	STRING	Name of POI category for audience interest measurement (e.g. Precisely standardized brand name used by the business)
7	BRAND_AFFINITY_CATEGORY_MICODES	STRING	Precisely POI classification reserved set of MiCodes[1] which was agreed to be the 1099**** Code space as the "Reserved Space"
8	BRAND_AFFINITY_SCORE_MINIMUM	FLOAT	Minimum Audience interest
9	BRAND_AFFINITY_SCORE_MAXIMUM	FLOAT	Maximum Audience interest
10	BRAND_AFFINITY_SCORE_MEAN	FLOAT	Mean Audience interest
11	BRAND_AFFINITY_SCORE_MEDIAN	FLOAT	Median Audience interest
12	BRAND AFFINITY_REACH	FLOAT	Target Reach

Notes:

[1] Click here to see the complete MiCode-to-SIC code lookup table

Footfall and Brand Affinity Parameters Equations

Glossary

Parameter is a number calculated according to some mathematically defined formula. It is derived from a geospatial dataset of **base signals** and it is tied to a geospatial **target**.

Signal is a geospatial record with mandatory latitude, longitude, userid and timestamp fields, and any number of optional fields (e.g., affinity_score, residency_score, geo-demogrsafics_value, movement_modality, visitation_modality, etc)**.**

User. The userid field of a **signal** defines a group of signals considered as belonging to the same **user**.

Audience is a set of **users** defined by some set of **constraints**.

Constraint is a particular rule to group and divide **users** into subsets. Constraints can be defined by a geospatial outline, date and time values, checks for an arbitrary numeric range, by string comparison, and others, depending on the needs.

Base outline is a large region defined by a geospatial boundary. It defines the **base** subset of **signals** that are a subject to calculate the formula.

Target outlines are smaller ones (that are usually contained within the **base** outline) define regions to calculate end values of the **parameter**.

Audience level is a particular set of such **constraints** for a **parameter** to reference in its formula. There can be as many levels defined for a particular **parameter**, but there is always is a **target** one.

Dwell-Time

$$mean\left(\left\{\frac{S}{S}\right\}_{a=1}^{A}\right) \times 100\%$$

where

Audience levels:

1. Base audience: entire area. Defined by a large outline (for example, United Kingdom). **BL**
2. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
3. General audience Level 1: dates in a period. **G1**
4. General audience Level 2: time for each period (for example, 08:00–21:59). **G2**

a – a user with at least one signal in **TA+G1+G2**

A – the **count** of all ***a*** users in **TA+G1+G2**

s – the **count** of all signals (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

S – the **count** of all signals of the user ***a*** in **BL+G1**

Average Number of Signals

$$\frac{\sum_{a=1}^{A} s}{A}$$

where

Audience levels:

1. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
2. General audience Level 1: dates in a period. **G1**
3. General audience Level 2: time for each period (for example, 08:00–21:59). **G2**

a – a user with at least one signal in **TA+G1+G2**

A – the **count** of all ***a*** users in **TA+G1+G2**

s – the **count** of all signals (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

Reach

$$\frac{A}{N_A} \times 100\%$$

where

Audience levels:

1. Base audience: entire area. Defined by a large outline (for example, United Kingdom). **BL**
2. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
3. General audience Level 1: dates in a period. **G1**
4. General audience Level 2: time for each day (for example, 08:00–21:59). **G2**

A – the **count** of all users in **TA+G1+G2**

$\boldsymbol{N_A}$ – the **total count** of all users in **BL+G1**

Proportion of Signals

$$\frac{\sum_{a=1}^{A} S}{\sum_{a=1}^{A} S} \times 100\%$$

where

Audience levels:

1. Base audience: entire area. Defined by a large outline (for example, United Kingdom). **BL**
2. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
3. General audience Level 1: dates in a period. **G1**
4. General audience Level 2: time for each period (for example, 08:00–21:59). **G2**

a – a user with at least one signal in **TA+G1+G2**

A – the **count** of all ***a*** users in **TA+G1+G2**

s – the **count** of all signals (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

S – the **count** of all signals of the user ***a*** in **BL+G1**

Footfall Score

$$\frac{A}{N_A} \times median\left(\left\{\left\{\frac{S}{S}\right\}_{j=1}^{s}\right\}_{a=1}^{A}\right)$$

where

Audience levels:

1. Base audience: entire area. Defined by a large outline (for example, United Kingdom). **BL**
2. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
3. General audience Level 1: dates in a period. **G1**
4. General audience Level 2: time for each period (for example, 08:00–21:59). **G2**

a – a user with at least one signal in **TA+G1+G2**

A – the **count** of all ***a*** users in **TA+G1+G2**

N_A – the **total count** of all users in **BL+G1**

j – a signal (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

s – the **count** of all signals (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

S – the **count** of all signals of the user ***a*** in **BL+G1**

Footfall Score Normalized

$$\frac{FOOTFALL\_SCORE}{FOOTFALL\_SCORE_{max}} \times 100\%$$

Footfall Score Rank

PERCENTRANK.INC function

Returns the rank of a value in a data set as a percentage (0..1, inclusive) of the data set.

Footfall Score SD

$$\frac{FOOTFALL\_SCORE - FOOTFALL\_SCORE_{mean}}{FOOTFALL\_SCORE_{standard\ Deviation}}$$

Brand Affinity Score

$$mean\left(\left\{\{Score_a\}_{j=1}^{s}\right\}_{a=1}^{A}\right)$$

where

Audience levels:

1. Target audience: defined by a designated hexagon, point with a radius, postcode, residence, and/or anything else. **TA**
2. General audience Level 1: dates in a period. **G1**
3. General audience Level 2: time for each period (for example, 08:00–21:59). **G2**

a – a user with at least one signal in **TA+G1+G2**

A – the **count** of all ***a*** users in **TA+G1+G2**

j – a signal (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

s – the **count** of all signals (latitude/longitude/time) of the user ***a*** in **TA+G1+G2**

$Score_a$ – Brand Affinity Score of the user ***a*** – the interest the selected category of physical locations (e.g., Points of Interest (POI) category) calculated according to the method described in the patent "Interest Profile of the User of the Mobile Application"

Brand Affinity Score Normalized

$$\frac{BRAND\_AFFINITY\_SCORE}{BRAND\_AFFINITY\_SCORE_{max}} \times 100\%$$

Brand Affinity Score Rank

PERCENTRANK.INC function

Returns the rank of a value in a data set as a percentage (0..1, inclusive) of the data set.

Brand Affinity Score SD

$$\frac{BRAND\_AFFINITY\_SCORE - BRAND\_AFFINITY\_SCORE_{mean}}{BRAND\_AFFINITY\_SCORE_{standard\ Deviation}}$$

Product Feedback and Support

Contact our Support team (info@locomizer.com) for product support and additional product information.

207 Regent Street
3rd Floor
London W1B 3HH
United Kingdom

www.locomizer.com