

World Premium Plus Points of Interest - Consumer Edition - USA & Canada

Version 5.0

UNITED STATES
www.pitneybowes.com/us
Technical Support:
<http://www.pitneybowes.com/us/support.html>

© 2014-2018 Pitney Bowes Software Inc.
All Rights Reserved.

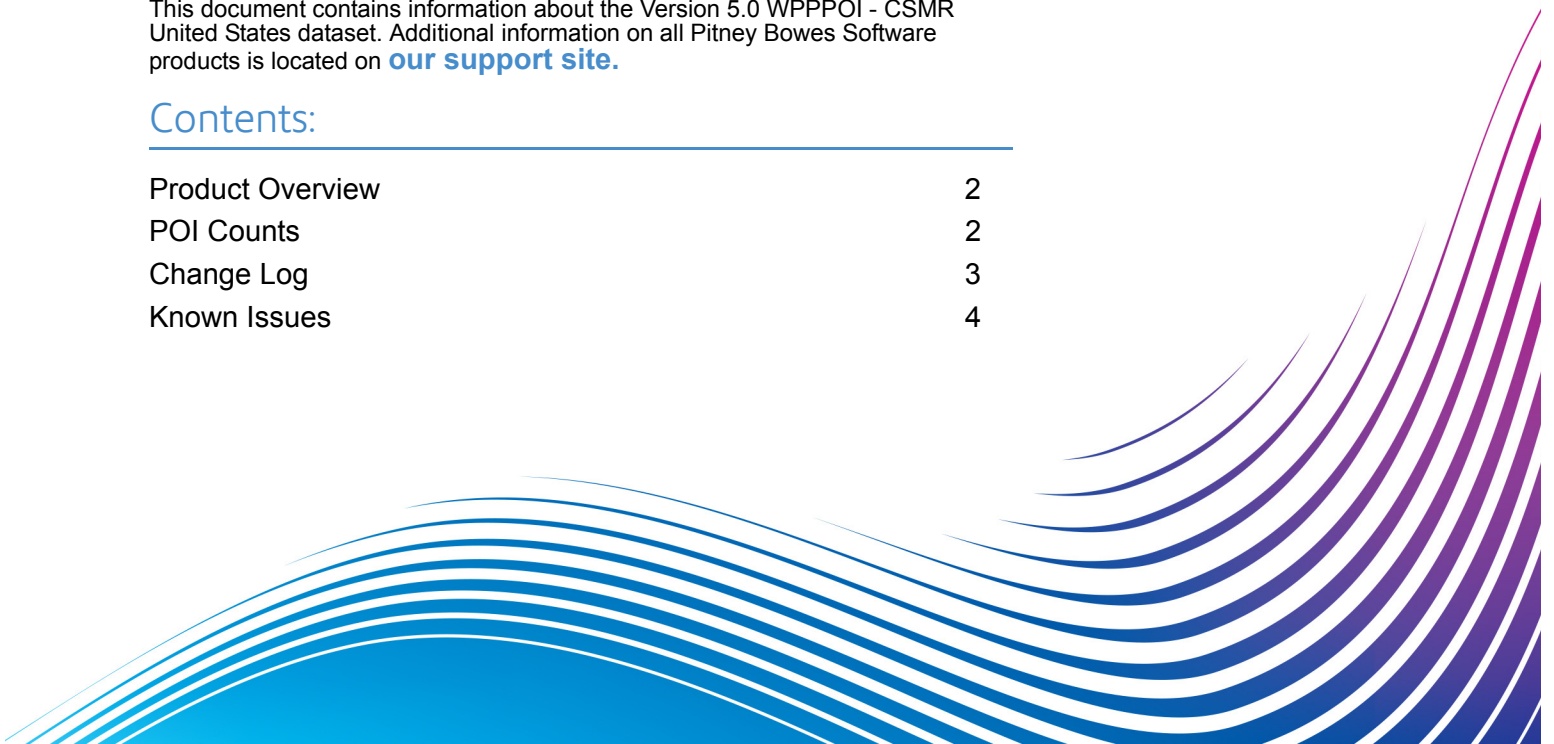
Release Notes

The World Premium Plus Points of Interest - Consumer Edition (WPPPOI - CSMR) United States Of America dataset is an innovative addition to our data portfolio. Pitney Bowes is committed to ongoing enhancements to this industry leading database to serve our customers' needs. Our road map is designed to continually maintain and improve the product so our customers can derive the highest value for their consumer services.

This document contains information about the Version 5.0 WPPPOI - CSMR United States dataset. Additional information on all Pitney Bowes Software products is located on [our support site](#).

Contents:

Product Overview	2
POI Counts	2
Change Log	3
Known Issues	4



Product Overview

The World Premium Plus Points of Interest – Consumer Edition (WPPPOI - CSMR) dataset provides information on the location of, and details relating to, a diverse set of consumer focused businesses, leisure hot spots and geographic features. The locational accuracy of this data has been further enhanced using Pitney Bowes Premium Spectrum Enterprise Geocoding solution with Master Location Data (MLD) to provide the most accurate location available for each Point of Interest.

Some of the feature of the WPPPOI - CSMR product are:

- Focused selection of POIs that concentrate on businesses, landmarks, and services relevant to consumers.
- Provides access to the largest USA and Canada geo-referenced POI dataset.
- The POIs are geocoded using Pitney Bowes comprehensive Global Geocoding capabilities with Ordnance Survey data.
- The POIs are classified using both the Pitney Bowes MiCode (providing a consumer-friendly classification across products) and the Standard Industrial Classification (SIC) codes.
- Please note that both these classifications are cross-referenced. The individual POI in the dataset have hierarchical relationships between themselves and others in their 'family'.
- The taxonomy ensures ease of use and global consistency.

The dataset, enables users to make informed decisions around access to services, retail or recreational facilities, location-based marketing and “find my nearest” searches.

POI Counts

Country	ISO3 Code	POI Counts
Canada	CAN	1,569,356
United States Of America	USA	16,778,392

Change Log

- Updated the Headers in schema to upper case
- A new column ALT_INDUSTRY_CODE has been added, after the SIC8_DESCRIPTION column in the schema
- Removed few columns such as, CEO_NAME and CEO_TITLE from the schema
- Position for GEORESULT and CONFIDENCE_CODE columns has been changed. Both these columns are moved after LATITUDE column in the schema.
- Changed the datatype for EMPLOYEE_HERE, EMPLOYEE_COUNT, YEAR_START and CURRENCY_CODE to Integer.
- Changed the datatype for SALES_VOLUME_LOCAL and SALES_VOLUME_US_DOLLARS to BigInteger
- Quality enhancements for multiple brands and new brandnames have also been added in USA. Click [here](#) for more details.
- Few brands have been updated with the new names. Click [here](#) for more details.
- False positives brandnames have been untagged
- Latitude and Longitude values may occasionally change slightly due to geocoding enhancements.
- Applied few changes in the address components for few records
- Removed few junk characters from the data (\, <, >, %, ^, ||)
- Improved the Brand Name standardization in the data and it has been standardized for respective countries:

Country	Brandname Counts
Canada	280
United States Of America	2,097

For detail view click [here](#).

Known Issues

There are some known issues with the source data. We will strive to address these issues in subsequent releases in a timely fashion. Please contact Technical Support with any questions you may have.

This section covers known issues and behaviors that require further explanation:

- Some duplicates may be encountered in the dataset.
- Some invalid values exist in the Name, areaName1, areaname3 and areaName, TradeName and Postcode column including instances where the Postcode column has values as 'zeroes'.
- Some diacritic loss in Name, TradeName, formattedAddress, mainAddressLine, addressLastLine, areaName1 and areaName3, CEO name, Domestic Ultimate Address, Global Ultimate Address and Parent Address columns.
- There are a few instances of text truncation exist for the Name, TradeName, Business Line, areaName3, formattedAddress, mainAddressLine, Http, Email and areaName2 columns.
- There are a few instances where the Employee Here (number of employees at the current location) count mismatches the Employee count (total number of employees).
- There are a few instances that have inconsistent Status Code and Subsidiary Indicator mapping.
- The Email, Http, Ticker symbol, and Exchange Name columns are not fully populated.
- The SIC8 (code) descriptions are undergoing updates. As a result, some values may not have any or updated description(s) and may have been removed.
- The administrative layer names are not standardized. As a result, some areaname4 values may be blank. There are a few instances where different variations of the same name are encountered in the areaName1 column.
- Mainaddressline and formattedaddress are not standardized in few records.
- There are a few instances where House number/Street Name is appearing twice in mainaddressline.
- There are a few instances where the areaname3 is populated with the name of a different admin level.
- Brandname standardization is work in progress, so there maybe some minor discrepancies.