



Navigating the Challenges of Data Quality and Recognition in China

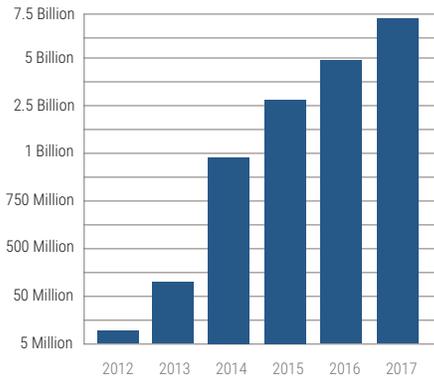
Global-Z China Experience

Building brand loyalty and a close relationship with the Chinese consumer is now a critical success factor for many premium international brands. The number of Chinese millionaires is expected to surpass that of any other nation by 2018, and by 2021 China is expected to have the most affluent households in the world.

A brand relationship is critical because “of the three brands considered for luxury purchase, two are considered top-of-mind, and are the ones that (Chinese) consumers actually buy on 93% of their purchasing occasions,” according to an August 2017 report from McKinsey & Company. “Wealthy Chinese are unlikely to purchase luxury goods that don’t fall within the two brands they consider top of mind.”

Global-Z International has been a major part of building the customer to brand relationship strategy for global premium brands for over twenty-five years and in the People’s Republic of China since 2003. Our address data quality expertise is built on a mature hygiene, matching and recognition process and the experience of processing 7.5-billion records with Chinese addresses. Today, it is one of Global-Z’s fastest growing markets, accounting for over 50% of our annual transaction volume.

Global-Z Shows Significant Growth in the Chinese Market



Cumulative cleansing transactions processed on Chinese name and address data.



Our extensive experience in China means that Global-Z can quickly and intelligently resolve the challenges related to data cleansing and matching Chinese customer records. For example, address formats are inconsistent. The Universal Postal Union, the Microsoft Developer Network reference on China addresses, and other sources have different preferred China address formats. Global-Z understands and can handle the differences.

Even more important than our domestic China experience, Global-Z can help to match Chinese consumers with records created abroad in other formats. A valuable part our experience is in identifying Chinese customers at home and overseas to provide the links needed to build a single customer view no matter where they are located.

“Two-thirds of Chinese luxury spending happens overseas,” the McKinsey report continues. Therefore, it is necessary for brands to bring together the purchase behavior of Chinese consumers both at home and abroad. The problem is that it is not easy to merge data that is built using different systems:

- Customer names entered in Chinese characters need to be transliterated to match consumer records with the names entered in other markets. However, transliteration systems give different results, making automatic matching difficult.
- Chinese names entered into Western systems are not always entered in the same way by data entry personnel. For example, the Chinese surnames Wang, Huang, and Wong all refer to the same surname.
- Address data is often inconsistent. As mentioned earlier, there are differences in preferred China address formats depending on sources. A Microsoft-based system using the Microsoft Developer Network format may not match other systems.
- Western systems often do not have all of the needed fields in the proper order to represent Chinese addresses. For example, in many languages, addresses are usually written from the smallest location to the largest In

China, this is the reverse. Here is a common way in which addresses would be written in the Mainland:

- 省 (shěng): province
- 县 (xiàn): county (you don't need the county for a large city)
- 市 (shì): city / town
- 区 (qū): district
- 大街 (dàjiē): avenue
- 路 (lù): road
- 楼 (lóu): building
- 室 (shì): flat / apartment (or 宅 for a house)

When Mr. Li is traveling abroad, he would want to provide more information than the fields provided in a typical Western data entry form. As a result, some address information may be combined or left out so that it could fit. The data quality would be poor and would need to be improved.

The challenge for creating a single customer view for a seamless experience, analysis, or forecasting is that data quality parsing and cleansing must be applied consistently for both the Chinese domestic data and the data captured overseas. A company with both Chinese market experience and broad international experience is required to make it work.

Parsing Example With Chinese Character Set

A Chinese individual may write his name and address in the following format:

266071山东省青岛市 香港东路6号5号楼8号室 李小方

The information that Mr. Li Xiaofang entered in the example above is as parsed and Romanized (transliterated) by the Global-Z system:

Given Name:	小方	Xiaofang
Surname:	李	Li
Room:	8	8
Building:	5	5
Number:	6	6
Direction:	东	Dong
Road:	香港	Xiang Gang
City:	青岛	Qunqdao
Province:	山东	Shandong
Postal Code:	266071	266071



Global-Z

v2/18