Data Quality Report Card

A look into your data quality

What is a Data Quality Report Card?

Experian's Data Quality Report Card is an analysis of your data set that discovers validity, irregularities and inconsistencies across your data set that could be impacting overall business and product performance. The Data Quality Report Card helps determine if further investigation is necessary and can help in remediation planning. In this snapshot, results shown are aggregated and provide high-level data quality observations.

What is the process to analyze my data?

There are four simple steps: Transfer the data to us through a secure system, we process the data using Experian's Aperture Data Studio™ platform, and we provide you the results in the report card. Aperture Data Studio is an industry-leading data quality platform that helps organizations identify data quality discrepancies. As part of the Data Quality Report Card, you'll get a:

1. Data profiling summary
2. Duplicate analysis
3. Address data analysis
4. Email data analysis
5. Phone data analysis

Top benefits of the Data Quality Report Card

Insight into the quality level of your data set

Identify existing data errors and issues

Understand target acquisition data

Discover potential duplicates hiding in your data set

Create a data-driven remediation plan

Data profiling summary

Data profiling provides insights into the completeness, uniqueness, and formatting of data fields, while not attesting to the accuracy of data. Data profiling and discovery helps organizations understand where there are opportunities to improve data sets and help to determine which actions can be taken.

Duplicate analysis

Utilizing a series of matching rules and fuzzy match logic, we can identify exact and partial matches across data sets. The duplicate analysis conducted on the data sets identifies exact matches, close, probable, possible, and unique relationships. Close, probable, possible matches may exist with familial relationships or arise when certain identifiers are exact, but others may be unique (same name, different address).

Address data analysis

Address information in the data set were checked against source of truth data sets, like the USPS, to determine if the formats are valid and if the addresses are deliverable. The Good full match category is ideal to ensure mail or packages reach their destination. Address validation can utilize parsing rules and work with available data provided on input to obtain a range of match codes. For example, if an apartment number is missing, there may be a multiple match scenario.

Email data analysis

In the email validation analysis, **we check your email list for typos and errors like missing punctuation or extra spaces, looking for spam traps, verifying domains can accept mail,** and testing each email address with an undetectable message to help check if the email address works

Phone data analysis

We can search data from nearly 1,500 telecommunications providers in over 200 countries worldwide, runs your list of numbers through our provider database, **checks if the number is active, and collects important information on the number—such as phone type, network status, and country of origin.**

Sample results: Data Quality Report Card

- 11% of records are missing email addresses
- 12% of records are missing mobile phone data
- Address data columns have missing information or incomplete addresses are being captured
- Over the amount of US states. Records require a standard format for state column.

- Many occurrences of close matches with different emails or phone numbers
- The biggest cluster size is made up of 11 records. This cluster has the same surname and address but different first name, email, and phone number
- 370 clusters have a cluster size greater than 2

5,213 addresses are unmatched, **at risk of return mail**

162,130 addresses are good full match, which is **ideal to ensure mail or packages reach their destination**

299 emails should be **used with caution,** 4% of which are undeliverable

8,690 emails are verified, and **you can send to them with confidence**

Top "email" account like none@none.com

6 phones are **other or invalid**

22 phones are **landline**

980 phones are **mobile**

Top mobile phone count in the data set like blank values or 999-999-9999