



## SUCCESS STORY

### « DATENKONSOLIDIERUNG UND BEREINIGUNG ZUR DATENMIGRATION »

- » *Erkennung von fehlerhaften Daten und Datenstrukturen*
- » *Konsolidierung, Bereinigung und Anreicherung mit Geodaten*
- » *Überführung der bereinigten Daten in ein normalisiertes Datenbankmodell zur Migration in ein Salesforce-CRM*

### Ausgangssituation

Das Unternehmen ist ein führender Asset Manager für Unternehmensimmobilien auf dem Europäischen Markt und somit der Branche Real Estate zuzuordnen. Das Unternehmen verfügt über ein CRM-System des Anbieters Salesforce sowie einige weitere IT-Systeme und viele relevante Excel-Dateien, z. B. über Makler-Objekte.

Das Immobilienunternehmen ist darum bemüht, nicht den Überblick über die regelmäßigen Angebote zu verlieren, da ansonsten Maklercourtage – oft in Millionenhöhe – eventuell mehrfach gezahlt werden müssen. Jedes Objekt und Makler-Angebot können dabei ganz verschiedene Attribute und Ausprägungen aufweisen. Das Unternehmen sah in den mangelhaft geführten Datentabellen von über 20 Jahren Immobilienangeboten ein enormes Potential für Analysen, als auch die frühzeitige Erkennung von bereits erhaltenen Angeboten.

### Maßnahmen – Methodik und Technologie

Zunächst wurde in sehr enger Zusammenarbeit mit dem Unternehmen eine Datenbereinigung durchgeführt. Hierbei war insbesondere ein tiefes Verständnis für die Geschäftslogik rund um die Daten von Nöten. Um dem Unternehmen stets die rückwirkende Nachvollziehbarkeit über die komplexen Bereinigungsmechanismen gewährleisten zu können, wurden alle wirkenden Arbeitsschritte stets dokumentiert.

Zur Schaffung einer bereinigten Datenbasis und somit einer Steigerung der Datenqualität wurden statistische Verfahren, insbesondere Ausreißeranalysen, sowie analytische Verfahren des Text Minings

bzw. des Natural Language Processing (NLP) eingesetzt. Die Programmiersprache Python unter Einbeziehung geeigneter Bibliotheken diente als Werkzeug zur Umsetzung der Datenbereinigung.

Im nächsten Schritt fand eine weitere Erhöhung der Datenqualität, durch Anreicherung mit Geodaten, statt. Hierfür wurden, unter Einsatz der Google Geocoding API, die Adressdaten validiert und falls notwendig rekonstruiert.

Im letzten Schritt wurden die konsolidierten, bereinigten und angereicherten Daten in ein normalisiertes Datenbankmodell in einem Microsoft SQL Server überführt. Dazu wurde zunächst ein ERM-Diagramm entworfen, welches der aktuellen sowie den zukünftigen Anforderungen und einer geplanten Migration nach Salesforce gerecht werden würde. Diesem Entwurf zugrundeliegend wurde die Datenbank aufgesetzt und dem Unternehmen zur Verfügung gestellt.

## Erfolge und Ergebnisse

Auf Basis dieser Fusion, Bereinigung und Validierung der historischen Daten konnten verlässliche Datenanalysen durchgeführt werden. Der generelle Mehrwert liegt in der Datenbankmodellierung, die sowohl eine Grundlage für anstehende Data Science Projekte bildet und zudem die Migration der Daten in das CRM-System ermöglicht. Das Datenmodell erlaubt keine redundanten Einträge mehr und Angebote können nur noch erschwert doppelt erfasst werden. Das Risiko von Doppel-Zahlungen der Maklercourtage wurde somit auf ein Minimum reduziert.

Darüber hinaus konnte erstmals ein umfassendes Reporting über die Angebote und bereits vorhandener Dopplungen erstellt werden. Für das Reporting kam das Tool Microsoft Power BI zum Einsatz.

## Was dies für Sie bedeutet

Daten sind die Entscheidungsgrundlage der heutigen Zeit! Ist die Qualität der Daten schlecht, so wird es auch die darauf basierende Entscheidung sein.

Setzen Sie auf Data Science und Data Engineering, können Sie die richtigen strategischen Entscheidungen treffen – und auf operativer Ebene sogar Entscheidungen und deren Ausführung automatisieren! Sie kommen dem daten-getriebenen Unternehmen damit einen entscheidenden Schritt näher und sichern Ihre Wettbewerbsfähigkeit.

Sie interessieren sich allgemein für die Möglichkeiten im Bereich der Vorhersage von Geschäftsereignissen und Finanzkennzahlen? Sie möchten mit uns eine individuelle Problemstellung besprechen? Besuchen Sie uns auf [www.datanomiq.de](http://www.datanomiq.de) oder schreiben Sie uns eine E-Mail an [info@datanomiq.de](mailto:info@datanomiq.de).



DATANOMIQ ist der herstellerunabhängige Beratungs- und Service-Partner für Business Intelligence und Data Science. Wir erschließen die vielfältigen Möglichkeiten durch Big Data und künstliche Intelligenz erstmalig in allen Bereichen der Wertschöpfungskette. Dabei setzen wir auf die besten Köpfe und das umfassendste Methoden- und Technologieportfolio für die Nutzung von Daten zur Geschäftsoptimierung.