

Case Study

Anomalieerkennung in Retail Daten

Zur Sicherstellung einer optimalen Datenqualität haben wir für unseren Retail Kunden ein Modell entwickelt, das ungewöhnliche Datenpunkte in Sales Daten automatisch erkennt und korrigiert.

STATWORX. *Data Science done right.*

STATWORX, Wöhlerstr. 8-10, 60323 Frankfurt
info@statworx.com, www.statworx.com

Herausforderung

Einwandfreie Daten stellen die Grundlage für alle Analytics-, Reporting- und Data Science Fragestellungen dar. Häufig müssen Daten vor der eigentlichen Auswertung aufwändig aufbereitet und bereinigt werden. Unser Kunde, ein internationales Retail Unternehmen, stand vor der Herausforderung, täglich eine vollautomatische Verifizierung und Bereinigung der Sales Daten angeschlossener Filialen durchführen zu müssen, um darauf aufbauende Reportingsysteme fehlerfrei bereitstellen zu können.

Ansatz

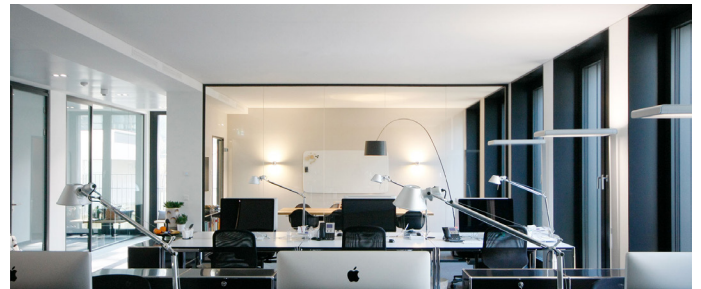
Basierend auf den Sales Zeitreihen der letzten zwei Jahre (ca. 500mio Datenpunkte) entwickelten wir, gemeinsam mit dem Kunden, ein statistisches Modell, das pro Produkt-Filial-Kombination die tatsächliche mit der empirisch beobachteten Verteilung der jeweiligen KPI abgleicht und so auffällige Datenpunkte vollautomatisch erkennen kann. Weiterhin ist das Modell in der Lage Anomalien in den Daten auf die durch das Modell erwarteten Werte zu glätten, um die jeweilige Beobachtung nicht vollständig löschen zu müssen. Der Algorithmus wurde vollständig in R entwickelt und auf dem vorhandenen Analytics Server in-Database deployed.

Ergebnis

Nach dem Deployment erkennt das Modell seitdem täglich Anomalien und auffällige Datenpunkte vollautomatisch. Durch die Anwendung des Modells ist es erstmalig gelungen, eine automatische Datenaufbereitung und Bereinigung der täglich angelieferten Filialdaten zu implementieren, die zuverlässige und stabile Ergebnisse liefert. Zudem fallen durch die Open Source Software R keine Lizenzkosten an.

Über STATWORX

STATWORX ist ein spezialisiertes Beratungsunternehmen für Data Science, Statistik und Machine Learning. Wir beraten, unterstützen und schulen Unternehmen bei der Entwicklung und Umsetzung von Data Science Projekten.



Unser interdisziplinäres Data Science Team besteht aus Mitarbeitern verschiedener akademischer Fachrichtungen wie z.B. Statistik, Informatik, BWL, VWL und Psychologie.

Heute ist STATWORX eines der wenigen Unternehmen im deutschsprachigen Raum, das *ausschließlich* Data Science Services anbietet. Seit der Gründung des Unternehmens im Jahr 2011 haben wir mehr als 100 Data Science und Analytics Projekte erfolgreich umgesetzt. *Data Science done right.*



Ihr Ansprechpartner

STATWORX

Fabian Müller

Leiter Data Science

fabian.mueller@statworx.com

+49 (0)69 6783 0675 - 1

www.statworx.com