

# Projekt "DataCo Supply Chain"

---

Michaela Nase, Matr. 1778563  
Frederik Schaefer, Matr. 1725833  
Michael Zemler, Matr. 1841490

12. Juni 2021

Datum der Präsentation .....12. Juni 2021  
Instruktor .....Dr. Florian Kellner

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung und Motivation</b>	<b>1</b>
1.1	Informationen über <i>DataCo Global</i> . . . . .	1
1.2	Intention . . . . .	1
<b>2</b>	<b>Aufbau und Ablauf der Applikation</b>	<b>2</b>
2.1	Datensatzbeschreibung und Datenbankanbindung . . . . .	2
2.2	Analytics . . . . .	3
2.2.1	Filter . . . . .	3
2.2.2	Kuchendiagramm- Bestellstatus . . . . .	3
2.2.3	Kuchendiagramm-Bezahlart . . . . .	4
2.2.4	Histogramm - Zeitverlauf der Bestellungen . . . . .	5
2.2.5	Tabelle - Top Kunden . . . . .	6
2.2.6	Tabelle - Top Items . . . . .	7
2.3	Kartenmodul . . . . .	8
2.3.1	Kunden . . . . .	8
2.3.2	Lieferungen . . . . .	10
2.4	Forecast . . . . .	12
<b>3</b>	<b>Anhang</b>	<b>14</b>
3.1	Übersicht der Projektdateien . . . . .	14

# 1 Einführung und Motivation

In diesem Projekt wurde eine Applikation zur Auswertung von Kunden- und Bestelldaten, basierend auf einem Datensatz der Firma *DataCo Global*, entwickelt. In der Dokumentation sollen zunächst kurz die Projektmotivation und Ideen zur Konzeption beschrieben werden, im Hauptteil folgt anschließend eine genaue Beschreibung der Applikation. Der Code der Applikation ist im Abgabeordner, alternativ kann sie über diesen Link (HIER Klicken!) betrachtet werden.

## 1.1 Informationen über *DataCo Global*

Die Firma *DataCo Global* ist im Bereich E-Commerce tätig. Kernaktivitäten sind Produktion, Verkauf sowie die kommerzielle Bereitstellung von Produkten. Das Unternehmen vertreibt Kleidung, Sportartikel und Elektronik. Die Optimierung der logistischen Prozesse steht hierbei stark im Fokus. Das wurde auch bei der Konzeption der Applikation berücksichtigt.

## 1.2 Intention

Online-Shopping ist mittlerweile ein fester Bestandteil des Einkaufserlebnisses vieler Kunden. Um sich gegen andere E-Commerce-Anbieter durchsetzen zu können, müssen Plattformen eine hohe Kundenzufriedenheit sicherstellen, da die Kunden andernfalls zur Konkurrenz abwandern. Anforderungen der Kunden an das Unternehmen wie beispielsweise kurze Lieferzeiten, bequemes Bezahlen oder Qualität der Produkte sollten im Fokus stehen.

Um im Blick zu behalten, wie es um die Performance von *DataCo Global* steht, soll eine Dashboard-Applikation entwickelt werden. Diese ermöglicht die Exploration und Analyse von bisher erfassten Kunden- und Bestelldaten. Des Weiteren sollen Funktionen implementiert werden, die es ermöglichen, Kundenpräferenzen wie z. B. beliebte Items oder Zahlungsmittel zu ermitteln. So kann Werbung gezielter erfolgen und die Kundenbindung gestärkt werden.

Mit Hilfe von Positionsdaten können die Standorte von Kunden visualisiert werden. Es können Cluster mit besonders hoher Kaufkraft identifiziert werden, die dann präferiert behandelt werden können. Außerdem kann basierend auf den Lieferzeiten der Standort für ein neues Logistikzentrum oder Ähnliches berechnet werden. Zusätzlich bieten Daten zu fehlgeschlagenen oder verspäteten Lieferungen eine Möglichkeit, frühzeitig Schwächen in der Lieferkette zu identifizieren und dann proaktiv darauf zu reagieren.

Die Analyse vergangener Verkaufszahlen kann helfen, zukünftige Verkäufe zu prognostizieren und somit Lagerbestände vorzuhalten. Auch das ermöglicht kürzere Lieferzeiten und stellt sicher, dass die Kunden alle nachgefragten Artikel jederzeit verfügbar haben. Nicht nachgefragte Produkte können aus dem Sortiment entfernt werden, um Lagerplatz zu sparen. Außerdem lassen sich Umsatzprognosen erstellen, um Liquidität vorherzusagen.

Die hier angesprochenen Wettbewerbsvorteile sollen durch die erstellte Applikation nutzbar werden. Im nächsten Kapitel werden die dafür entwickelten Module genauer beschrieben.

## 2 Aufbau und Ablauf der Applikation

Die fertige Applikation gliedert sich in drei Module.

Im Analysemodul, werden auf einen Blick Informationen wie Status der Bestellungen oder beliebige Items ersichtlich. Mit Hilfe von Filtern können die Informationen angepasst werden.

Im Kartenmodul gibt es zwei verschiedene Ansichten. In der ersten Ansicht wird die Position der Kunden auf der Weltkarte angezeigt. Es soll so möglich werden, Gebiete mit einer hohen Zahl profitabler Kunden ausfindig zu machen. Basierend darauf kann ein neuer Lagerstandort errechnet werden. In der zweiten Ansicht werden Informationen über aktuelle und fehlgeschlagene Bestellungen auf der Karte visualisiert. Hier soll es möglich sein, Probleme in der Lieferkette zu entdecken. In beiden Ansichten werden relevante Informationen unter der Karte angezeigt.

Im Forecastmodul soll eine erste Umsatzplanung ermöglicht werden. Der Fokus liegt hier vor allem auf einer umfassenden Filtermöglichkeit der verkauften Produkte und einer möglichst flexiblen Gestaltung des Vorhersagehorizontes. Basierend auf den Ergebnissen der Umsatzplanung soll so eine bessere Kalkulation der Lagerbestände erzielt werden. Ebenfalls erhofft man sich dadurch, dem Umsatz profitabler Produkte isoliert betrachtet besser kalkulieren zu können.

### 2.1 Datensatzbeschreibung und Datenbankbindung

Die Daten für die Applikation wurden auf kaggle.com im CSV-Format bereitgestellt. Konkret handelt es sich dabei um einen Datensatz von *DataCo Global* mit Informationen zu Kunden und Bestellungen in 53 Spalten. Diese enthalten Informationen über Kunden, Liefermodalitäten sowie Produkte. Es werden beispielsweise Angaben über Zahlungsart, Lieferzeiten, Gewinne, Produktpreise oder Adresse der Kunden gemacht.

Die CSV-Datei wurde in eine SQLite-Datenbank umgewandelt. Anschließend wurden noch kleinere Änderungen an den Daten durchgeführt um diese besser nutzbar zu machen. Da beim Bestelldatum auch die Uhrzeit der Bestellung angegeben war, die für Auswertungen aber nicht von Bedeutung ist, wurde das Datum bereinigt. Die spanische Bezeichnung *EE*. *UU*. wurde durch *USA* ersetzt.

## 2.2 Analytics

In der Sektion *Analytics* können die Bestellungen ausgewertet werden.

### 2.2.1 Filter

Filters

Category (50)	Customertype (3)	Market (5)	Department (11)
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Abbildung 1: Filter

Mit Hilfe von Filtern können die dargestellten Grafiken nach Produktkategorien, Kundentypen, Märkten oder Abteilungen durchsucht werden. Dies kann hilfreich sein, wenn man wissen möchte, wie die gewählten Entitäten performen. Eine Unterscheidung kann wichtig sein, da zum Beispiel Kunden unterschiedlicher Departments unterschiedliche Erwartungen an Produkte oder Preise haben. Es kann außerdem sein, dass einzelne Kategorien besser oder schlechter als der Rest performen. Kennt man diese, kann man gute Kategorien ausbauen oder schlechte verbessern.

### 2.2.2 Kuchendiagramm- Bestellstatus

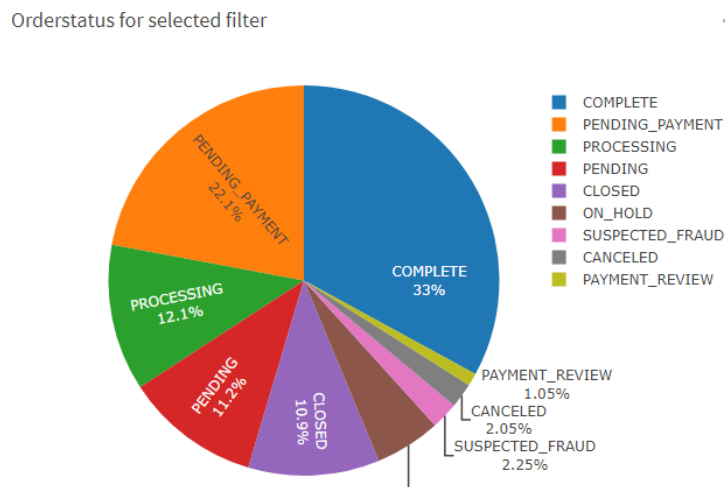


Abbildung 2: Kuchendiagramm - Bestellstatus

Ein Kuchendiagramm stellt den Status aller Bestellungen dar. So ist ersichtlich, ob viele offene Bestellungen vorliegen, es häufig zu Problemen bei der Bezahlung kommt oder ein Betrug vermutet wird. Des weiteren wird ersichtlich, ob viele Bestellungen abgebrochen werden.

### Mögliche Anwendungen:

Es ist auf einen Blick ersichtlich, welcher Anteil der Bestellungen abgeschlossen ist und in welchem Zustand sich alle anderen Bestellungen befinden. Fällt zum Beispiel auf, dass viele Aufträge abgebrochen werden sollte nachgeforscht werden, warum dies der Fall ist (Kommt die Ware eventuell beschädigt beim Kunden an? Gibt es dasselbe Produkt woanders günstiger?). Wichtig ist auch, die Rate an vermuteten Betrugsversuchen im Auge zu behalten. Nimmt diese drastisch zu, besteht eventuell eine Sicherheitslücke.

### 2.2.3 Kuchendiagramm-Bezahlart

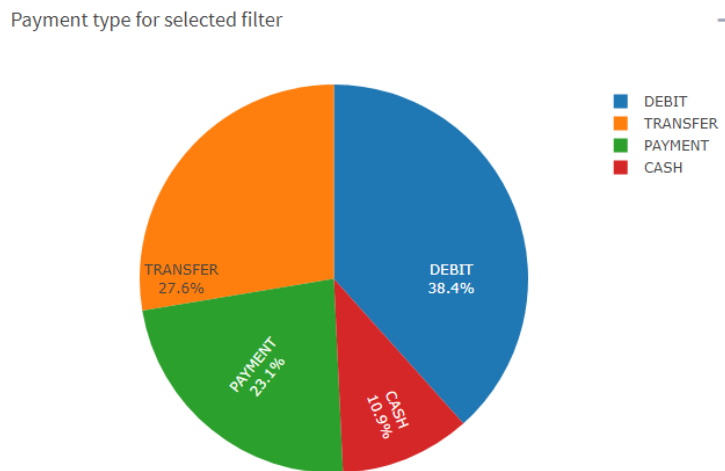


Abbildung 3: Kuchendiagramm-Bezahlart

Ein zweites Kuchendiagramm zeigt, wie die verschiedenen angebotenen Zahlungsarten (Debit, Payment, Transfer, Cash) angenommen werden.

### Mögliche Anwendungen:

Für die am häufigsten genutzten Zahlarten sollte sichergestellt werden, dass sie immer funktionieren. So sollten zum Beispiel bei Updates die Schnittstellen für am häufigsten genutzte Zahlverfahren oberste Priorität haben. Fällt auf, dass eine Bezahlart sehr wenig genutzt wird kann überlegt werden ob es sich lohnt, diese überhaupt noch anzubieten.

## 2.2.4 Histogramm - Zeitverlauf der Bestellungen

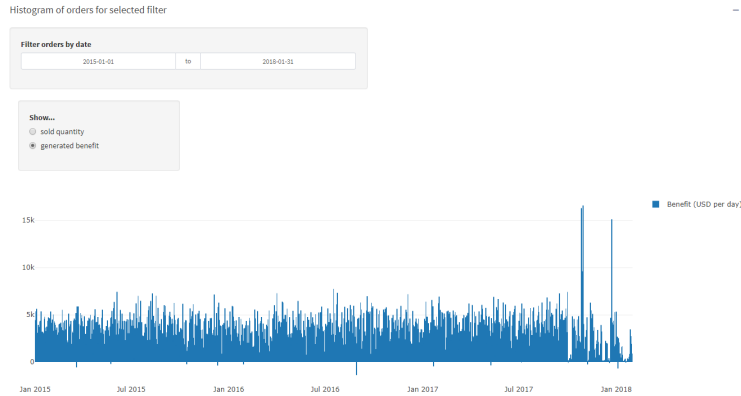


Abbildung 4: Histogramm der Bestellungen

Mit Hilfe eines Histogramms kann die zeitliche Entwicklung von entweder Bestellmenge oder Gewinn verfolgt werden. Mit Hilfe von Radio Buttons kann zwischen den beiden Darstellungen gewechselt werden. Will man nur ein bestimmtes Zeitfenster betrachten, so kann dies mit Hilfe eines Date Range Inputs gewählt werden.

### Mögliche Anwendungen:

Ein Einbrechen des Gewinns/der Bestellmenge ist auf einen Blick ersichtlich. Mit Hilfe der Datumswahl kann das betroffene Zeitfenster näher betrachtet werden. So kann schnell ermittelt werden, wann die Menge an Bestellungen bzw. der Gewinn zurrückging. Anschließend kann man sich Gedanken machen, was der Auslöser dafür gewesen sein könnte. Gab es eventuell Lieferschwierigkeiten, die Kunden verärgert haben könnten? Oder hat die Konkurrenz zu dieser Zeit verstärkt Werbung gemacht?

Außerdem kann verglichen werden, ob sich sowohl die Bestellmenge als auch der Gewinn verändert haben oder nur eines von beidem. Bleibt zum Beispiel der Gewinn gleich aber die Bestellmenge fällt, so ist das grundsätzlich kein Grund etwas zu unternehmen, es bedeutet eventuell nur, dass teurere, dafür weniger Items bestellt werden. Fällt hingegen der Gewinn bei gleich bleibender Bestellmenge, so werden billigere Items gekauft und man könnte versuchen, eine etwas teurere Premiumvariante der am häufigsten gekauften Items anzubieten.

## 2.2.5 Tabelle - Top Kunden

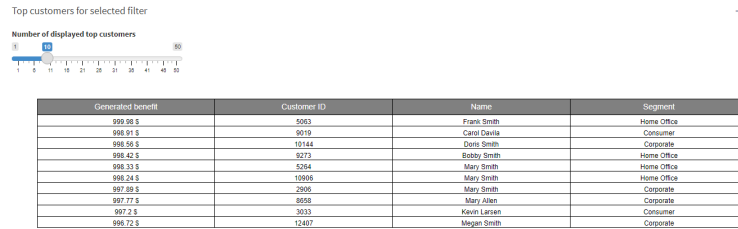


Abbildung 5: Tabelle der Top Kunden

Die Top Kunden des Unternehmens werden in einer Tabelle dargestellt. Dabei werden absteigend diejenigen Kunden angezeigt, die *DataCo Global* den höchsten Gewinn eingebracht haben. Neben dem Kundennamen und dem generierten Gewinn wird außerdem die Kundenart und die KundenID aufgelistet. Wie viele Top Kunden man sich anzeigen lassen will, kann mit Hilfe eines Sliders eingestellt werden.

### Mögliche Anwendungen:

Top-Kunden sind wertvoll für das Unternehmen, da sie den höchsten Gewinn generieren. Daher sollte auf jeden Fall versucht werden, die Kundenbindung für diese Personen sicherzustellen. Beschwerden der Top-Kunden sollten ernst genommen werden und es kann sich lohnen, ihnen bei Nachfrage etwas mehr Rabatt zu geben. Persönliche Werbung wie Anrufe oder Emails sind zeit- und kostenaufwändig und sollten daher auf die Top Kunden beschränkt werden, da es sich lohnen kann, wenn diese dem Unternehmen dadurch erhalten bleiben.



## 2.2.6 Tabelle - Top Items

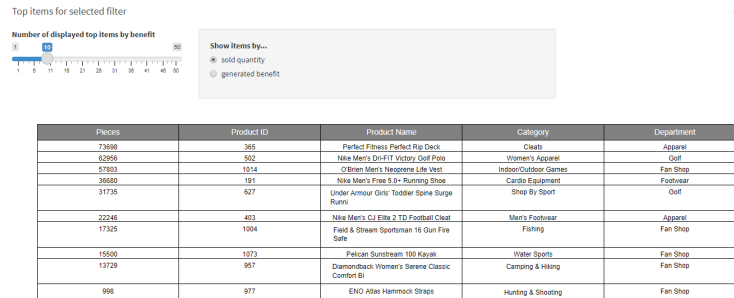


Abbildung 6: Tabelle der Top Items

Eine weitere Tabelle listet absteigend die Top Items von DataCo auf. Diese können entweder nach verkaufter Stückzahl oder nach erzieltm Gewinn angezeigt werden. Dies kann mit Hilfe von Radio Buttons gewählt werden. Neben der Stückzahl bzw. dem Gewinn werden die ProduktID, der Produktname, die Produktkategorie und die zugehörige Abteilung angezeigt. Wie viele Top Items man angezeigt bekommen möchte ist, wie auch schon bei den Top Kunden, mit einem Slider wählbar.

### Mögliche Anwendungen:

Items, die oft verkauft werden landen bei sehr vielen Kunden des Unternehmens, sie sollten daher von guter Qualität sein, damit Kunden ein gutes Bild vom Unternehmen haben. Es kann auch sinnvoll sein zu analysieren, was die Top Items gemeinsam haben und sich den Grund ihrer Beliebtheit zu überlegen, denn hier könnten Stärken des Unternehmens liegen. Diese können mehr in den Fokus gestellt werden. Die Itempalette könnte beispielsweise um ähnliche Items erweitert werden, damit die Kunden auch in Zukunft etwas beim Unternehmen bestellen.

Items, die hohe Gewinne erzielen sind ebenfalls wertvoll für das Unternehmen. Es kann versucht werden, durch Werbeaktionen mehr von diesen Items zu verkaufen. Denkbar wäre hier eine Rabattaktion wie zum Beispiel zwei zum Preis von einem oder verstärktes aufmerksam machen auf das Item beispielsweise auf der Website, im Newsletter oder als "das könnte Sie auch interessieren Anzeige nach einem Einkauf.

## 2.3 Kartenmodul

### 2.3.1 Kunden

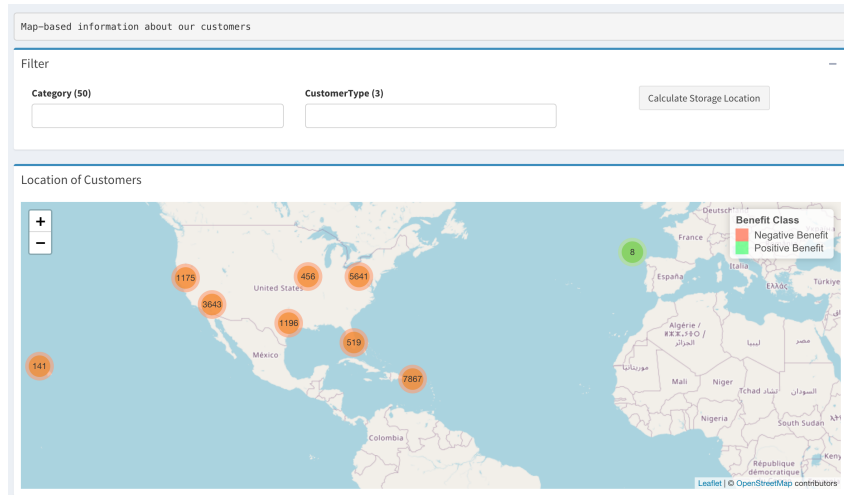


Abbildung 7: Karte mit Kundenposition

Auf der ersten Karte kann die geografische Position der Kunden untersucht werden. Die Kunden werden als Marker auf der Karte angezeigt. Ist der Kunde für das Unternehmen profitabel (basierend auf den Verkäufen), wird er grün markiert. Ansonsten wird er rot markiert. Die Kundenmarker werden geclustert, um bei niedrigem Zoom-Level eine gute Performance und Übersichtlichkeit der Karte zu gewährleisten. Die Marker auf der Karte sind interaktiv und zeigen bei einem Klick darauf weitere Daten über einen einzelnen Kunden an (Abbildung 8).

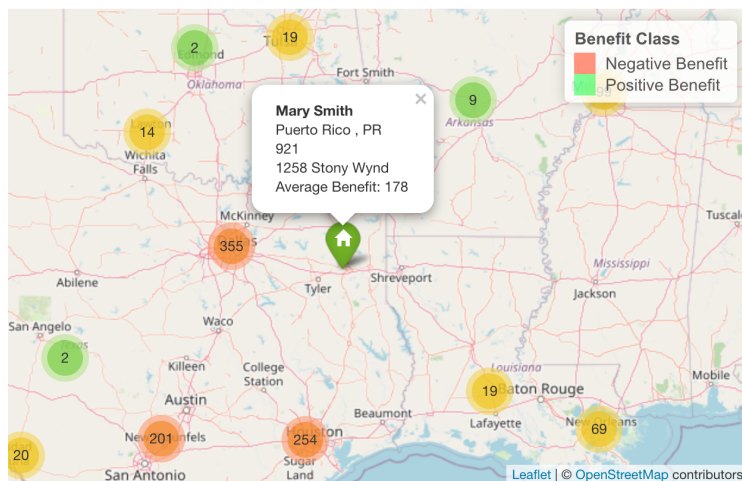


Abbildung 8: Popups und Legende

Oben im Modul können Filter zur Produkt-Kategorie und Kundenart eingestellt werden. Die auf der Karte angezeigten Marker werden entsprechend den ausgewählten Filtern angepasst. Zusätzlich befinden sich unter der Karte noch vier Infoboxen mit Statistiken, die sich auch je nach Filter anpassen (Abbildung 9). Des Weiteren gibt es einen Action-Button, der die Berechnung eines neuen Lagerstandortes durchführt. Dafür werden die Geo-Positionen aller gefilterten Kunden in Betracht gezogen und nach benötigter Lieferdauer gewichtet. Je höher die durchschnittliche Lieferdauer zu einem Kunden, desto größer seine Gewichtung bei der Berechnung. Ziel ist es, die Lieferzeit zu minimieren. Es wird die *Center of Gravity Methode* benutzt, um den Lagerstandort zu kalkulieren. Dieser wird dann mit einem blauen Marker und einem Popup mit den Koordinaten auf der Karte eingetragen (Abbildung 10).





	CUSTOMER COUNT <b>20652</b>		TOTAL ORDERS <b>180519</b>
	AVERAGE PRODUCT PRICE <b>141.23</b>		TOTAL BENEFIT <b>3966902.97</b>

Abbildung 9: Kundenstatistiken

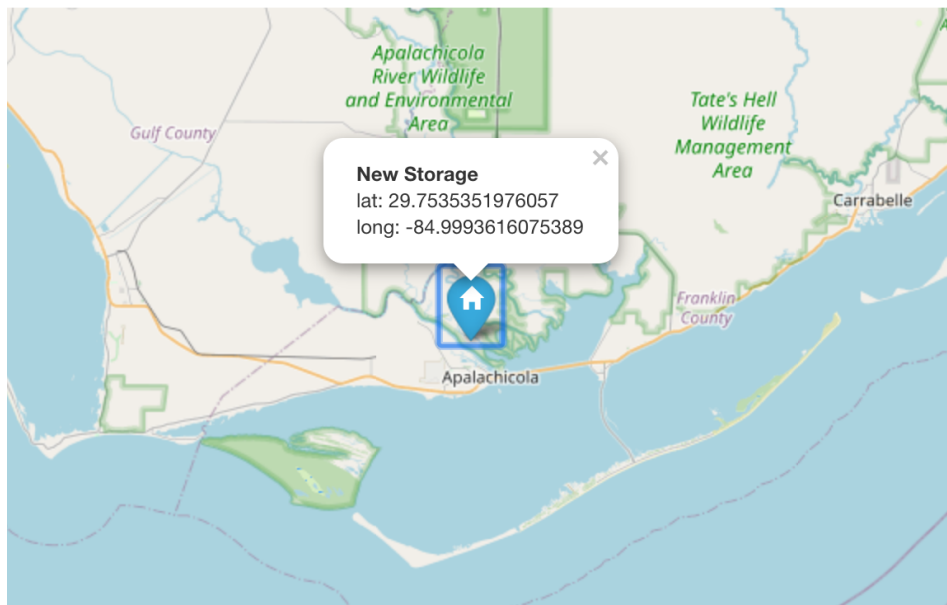


Abbildung 10: Neuer Lagerstandort

### 2.3.2 Lieferungen

Auf der zweiten Karte können die Lieferungen untersucht werden. Einzelne Lieferungen werden als Kreise auf der Karte angezeigt (Abbildung 11). War die Lieferung pünktlich, ist der Kreis grün, ansonsten wird er rot dargestellt. So können schnell Regionen auf der Karte identifiziert werden, bei denen es öfter zu Problemen kommt. Auch hier wird wieder Marker-Clustering verwendet, um die Performance der Karte bei hoher Markerzahl zu gewährleisten.

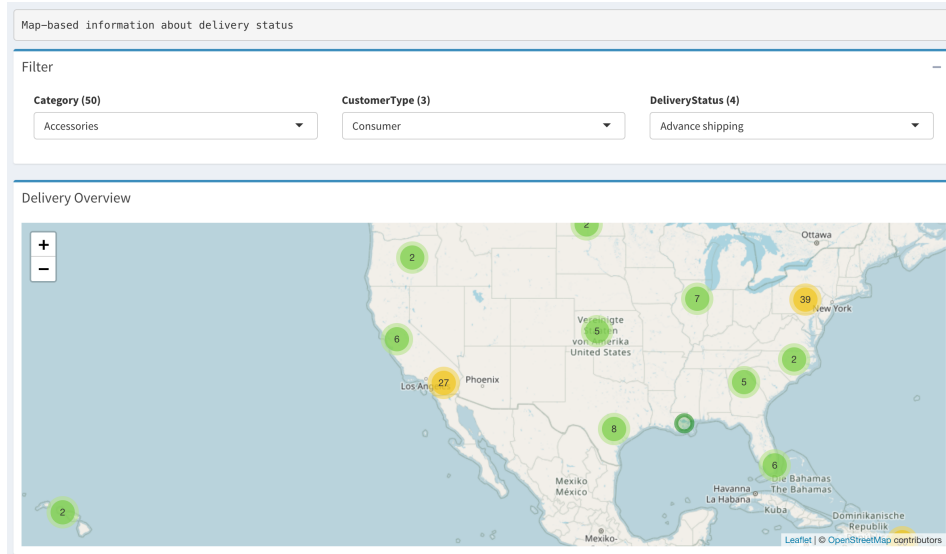


Abbildung 11: Lieferungen auf der Karte

Wie auch bei der Kundensicht, kann man mit einem Klick auf einen Marker weitere Informationen über die spezifische Lieferung erhalten. Ein Popup mit Daten zu der Lieferung wird angezeigt (Abbildung 12). Besonders interessant sind hier die Werte *Days estimated* und *Days actual*. Basierend auf diesen Werten wird ausgerechnet, ob es zu einer Lieferverzögerung kam.

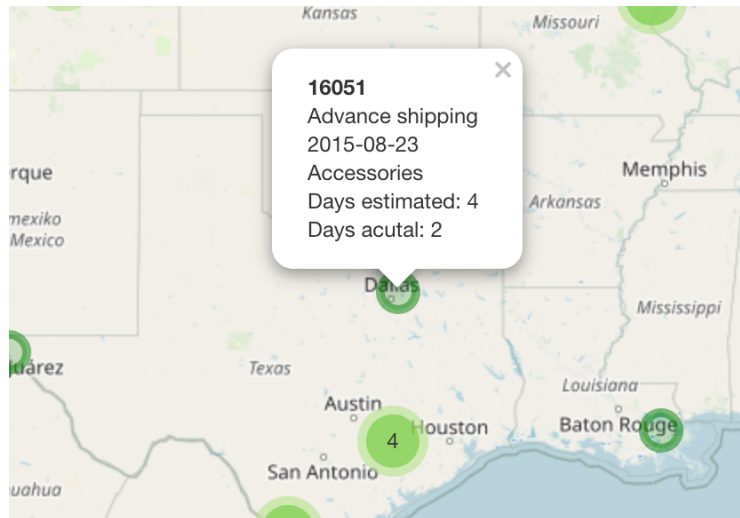


Abbildung 12: Popup mit Lieferungsinformationen

Es ist auch möglich, nach *Produktkategorie*, *Kundentyp* und *Lieferstatus* zu filtern. Die Karte reagiert dann auf die gesetzten Filter. Zusätzlich werden noch Informationen (Abbildung 13) zu den ausgewählten Lieferungen unter der Karte in Infoboxen angezeigt. Führt eine Filterkombination dazu, dass keine Daten angezeigt werden können, erscheint unten rechts eine kleine Warnung (Abbildung 14). Diese Warnungen wurden in der gesamten Applikation verbaut.

	TOTAL ORDERS <b>901</b>		DELIVERY ON TIME <b>152</b>
	PENDING <b>202</b>		DELIVERY DELAYED / CANCELED <b>547</b>

Abbildung 13: Informationen zu den Lieferungen

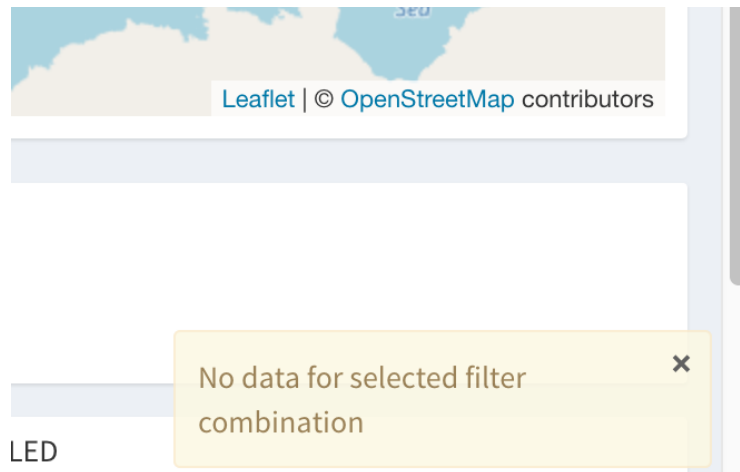


Abbildung 14: Fehlermeldung bei leerer Filterauswahl

## 2.4 Forecast

Es wurde eine Umsatz-Vorhersagefunktion implementiert, die *DataCo Global* nutzen könnte, um den Lagerbestand genauer zu planen: Wird so für die kommende Zeit ein höherer Umsatz für Produkt X vorhergesagt, kann dieses Produkt gehäuft gelagert werden, um Lieferengpässe zu vermeiden. Des Weiteren können mit gezielten Vorhersagen eben jene Produkte identifiziert werden, die einen Abwärtstrend bei den Verkaufszahlen zeigen und so eventuell aus dem Sortiment genommen werden sollten.

Um die benötigten Vorhersagen durchführen zu können, bietet dieses Modul zum einen mehrere Filter, die Produkte nach Absatzregion oder Kategorie filtern können, zum anderen kann der Vorhersagehorizont beliebig gewählt werden.

 A screenshot of a web form titled "Forecast of Sales". The form is divided into two main sections. The first section, "Forecast Horizon", contains a "Time Units" dropdown menu set to "Days" and an "Amount" input field with the value "10". The second section, "Apply Filters", contains three input fields: "Category (50)", "CustomerType (3)", and "Region (23)". Each input field has a small number in parentheses next to it, indicating the number of items available for selection.

Abbildung 15: Filtereinstellungen der Prognose

Im Hintergrund der Vorhersagefunktion arbeiten ein Simple-Exponential-Smoothing Algorithmus und ein Neuronales Netzwerk. Deren Ergebnisse fließen mit je 50 Prozent in die finale Vorhersagegewichtung ein. Der Grundgedanke hierbei ist, mit dem Simple-Exponential-Smoothing eine leicht zu verstehende Ausgangsbasis für eine Vorhersage zu erhalten. Diese wird mit der Absicht, tieferes Wissen aufzudecken, durch das Neuronale Netz erweitert. Abbildung 16 zeigt beispielhaft eine so erzeugte Vorhersage der nächsten vier Wochen.

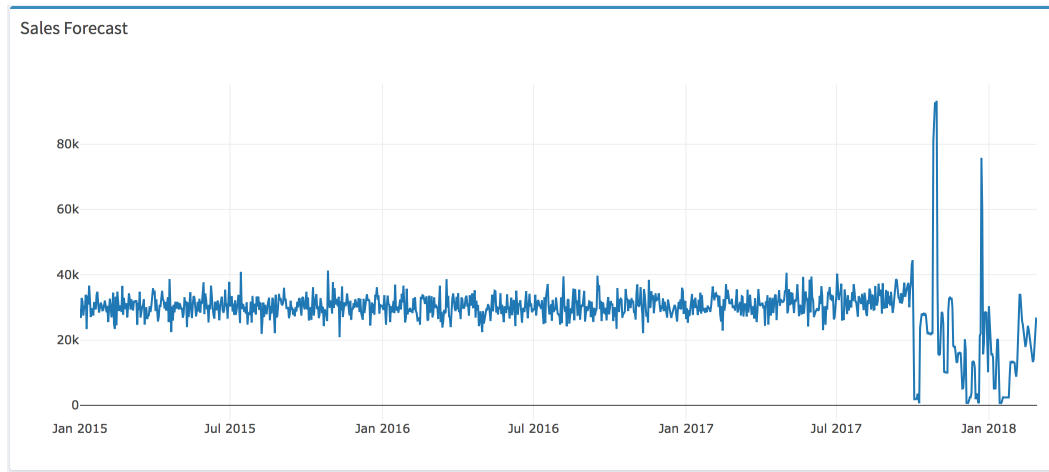


Abbildung 16: Beispielprognose des nächsten Monats über alle Verkäufe

Unterhalb der Prognose befinden sich zwei zusätzliche Kennzahlen: der erwartete Gesamtumsatz für den gewählten Prognosezeitraum und der durchschnittliche Umsatz pro Tag über die vorhergesagte Zeitspanne. Die zweite Kennzahl ist wichtig, da der Umsatz von Tag zu Tag schwanken kann.

## 3 Anhang

### 3.1 Übersicht der Projektdateien

	Dateiname	Beschreibung
	app.R	Hier wird das Shiny-Dashboard aufgesetzt und die Module in Tabs geladen
	data_import.R	Hier werden die Daten aus der Datenbank geladen
	DataCoSupplyChainDataset.db	Das ist die verwendete Datenbank. Muss im Projektordner liegen
	global.R	Hier werden die Pfade zu Modulen und der Library-File spezifiziert
	libraries.R	Hier werden die benötigten Pakete geladen/installiert. Pacman automatisiert den Prozess der Installation von Paketen und muss deshalb installiert werden!
M O D U L E	Analytics.R	Modul mit Daten und Statistiken rund um den Datensatz. Wird über global.R eingebunden und in app.R in einen Dashboard-Tab geladen
	ForecastView.R	Modul zur Prognose von Verkaufszahlen Wird über global.R eingebunden und in app.R in einen Dashboard-Tab geladen
	MapViewCustomer.R	Modul für kartenbasierte Informationen zu Kunden Wird über global.R eingebunden und in app.R in einen Dashboard-Tab geladen
	MapViewDelivery.R	Modul für kartenbasierte Informationen zu Lieferungen Wird über global.R eingebunden und in app.R in einen Dashboard-Tab geladen