

ReComMeND

Prof. Dr. Robert Knappe, HWR Berlin

Jonas Krembsler, HWR Berlin

Controlling Inspiration Berlin / 7. November 2020

Projektpartner:



Lufthansa Systems



**Lufthansa
Industry Solutions**

Gefördert durch:



Ertragscontrolling mit mathematischer Modellierung im öffentlichen Personennahverkehr am Beispiel der Berliner Verkehrsbetriebe AöR

Agenda:

1. Ertragsprognosen der BVG für Verkehrsdienstleistungen
2. Das Forschungsprojekt ReComMeND
3. Controlling-Kontext
4. Erste Ergebnisse
5. Ausblick und Einladung

1. Ertragsprognosen der BVG für Verkehrsdienstleistungen





3 Monate umsonst.

Kennen viele vom Praktikum.

Quelle: BVG

Imagevideo:

https://youtu.be/1Pipy_7nyr0

07.11.2020



**Als öffentliches Unternehmen dürfen wir
nicht verraten, wem wir den Wahlsieg gönnen.**

Deswegen nur so viel:

WIR WÜNSCHEN'S BEIDEN.



Ringo Starr

spielt heute in Berlin:

**Endlich ist jemand in der Stadt, der noch
schlechter den Takt halten kann.**

BVG



Kurzer Blick in den Spiegel:

Wir haben die Zahlen schön.

Überblick



Flächengröße Berlins



Abfahrten pro Tag



Unternehmensbezogene
Fahrgastfahrten (UBF)



Abonnenten



Einwohner

Fahrgastfahrten



U-Bahn



Straßenbahn



Omnibus Inkl. Föhre



Föhre

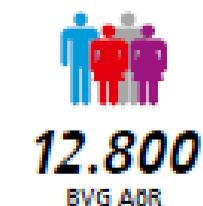


Fahrten BerKönig



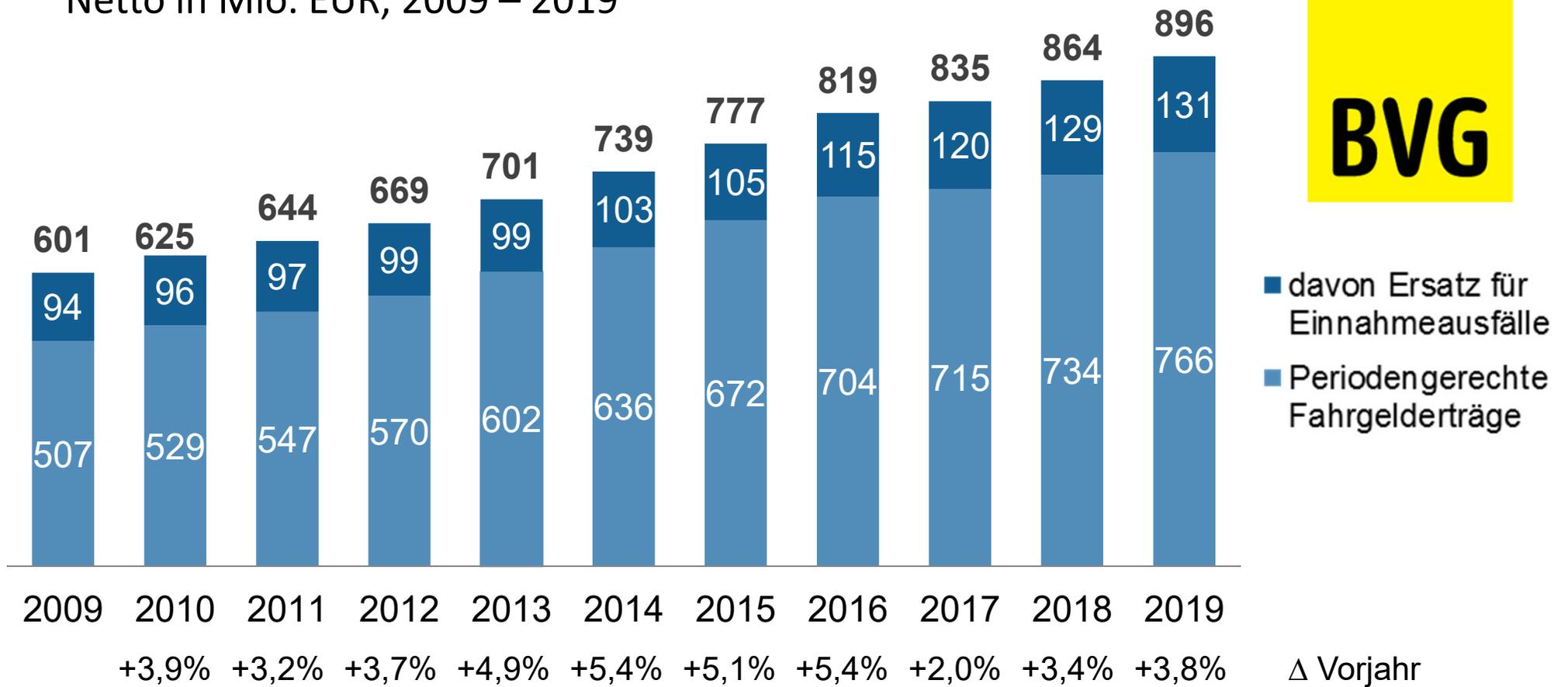
Zentraler Omnibusbahnhof (IOB)
jährliche An- und Abfahrten

Beschäftigte



Verkehrserträge BVG

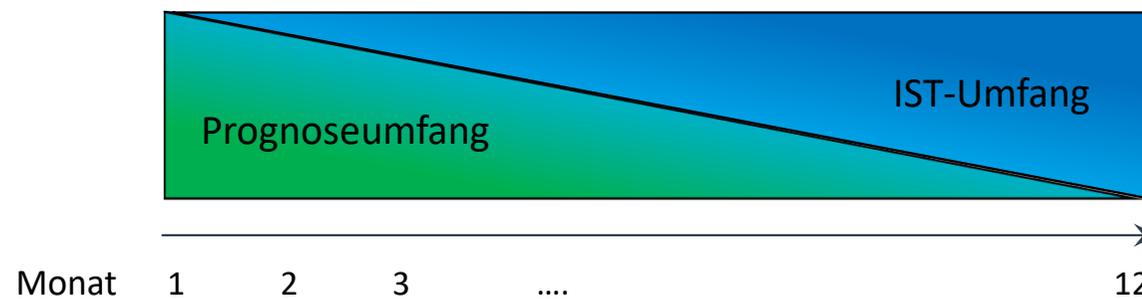
Netto in Mio. EUR, 2009 – 2019



Status Quo Ertragsprognose

Chronologie der monatsbasierten rollierenden Prognose:

11x per anno monatsweise Ersatz V-IST durch IST und Aktualisierung V-IST bis Dez.

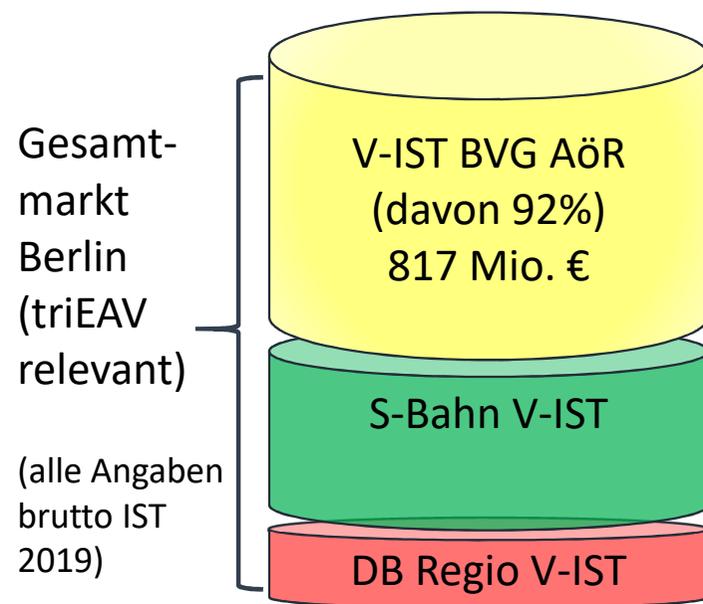


PLAN (statisch, Mitte Vorjahr) → V-IST auf Monatsbasis immer bis Dez. → IST

Kontext Fahrgelderträge triEAV

triEAV = trilateraler
Einnahmeaufteilungsvertrag

V-IST isoliert in 3 Unternehmen

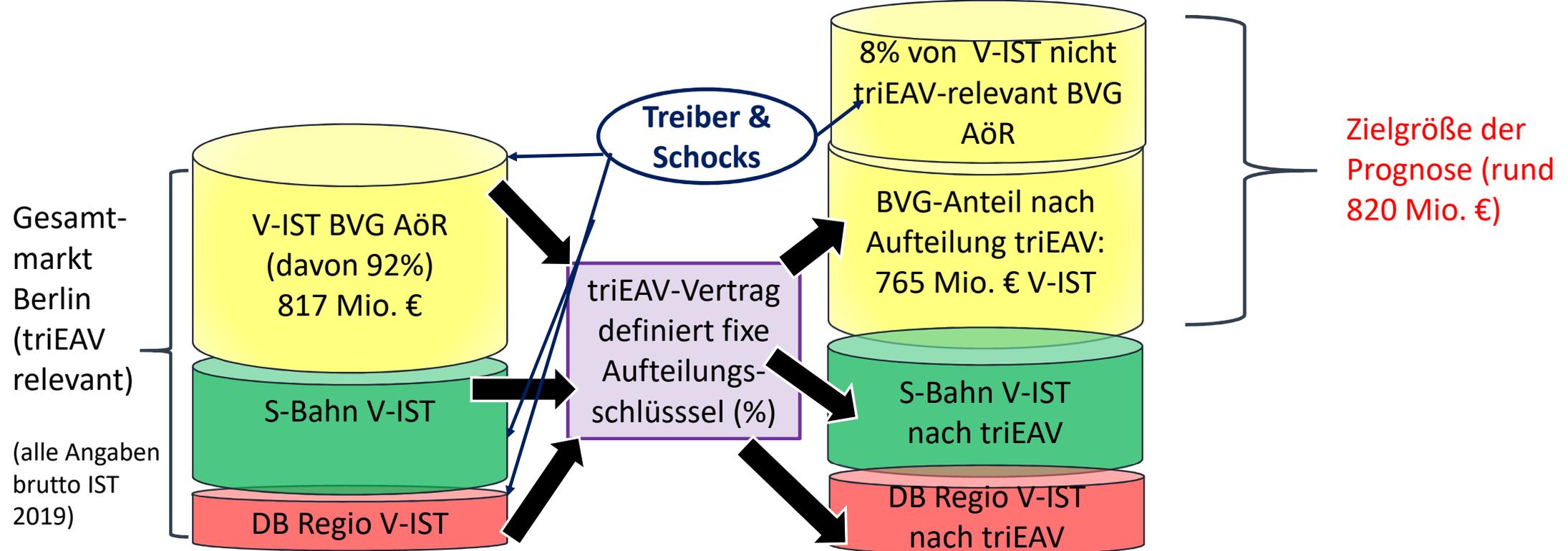


Kontext Fahrgelderträge triEAV

triEAV = trilateraler
Einnahmeaufteilungsvertrag

V-IST isoliert in 3 Unternehmen

V-IST nach
Einnahmeumverteilung



Bislang verwendete Treiber der BVG

Externe Treiber:

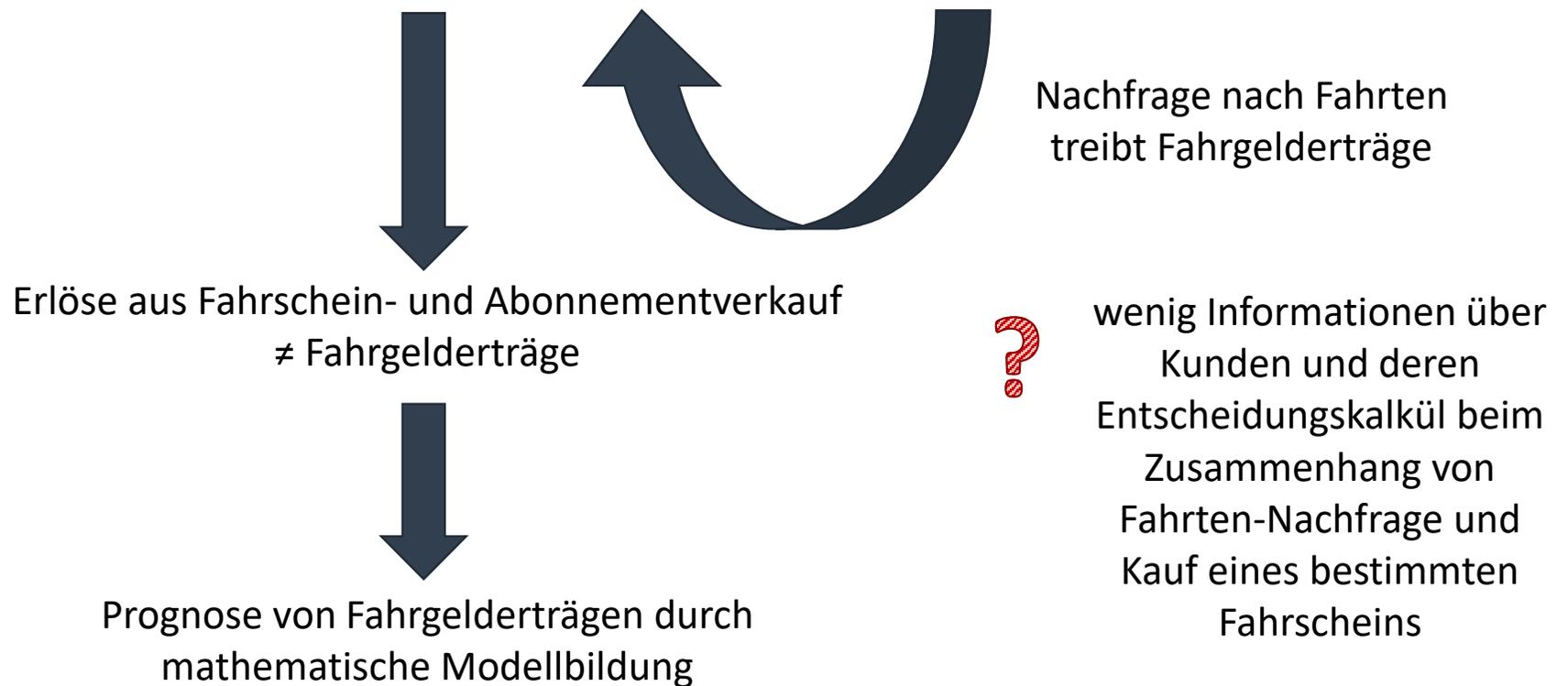
- Touristen in Berlin
- Pendlerbewegungen Umland
- Bevölkerungsentwicklung Berlin
- Kalendereffekte

Interne Treiber:

- Erfolg der Abokampagne
- Tarifmaßnahmen
(Preisgestaltung)

Betriebswirtschaftliche Logik

Fahrgelderlöse \neq Fahrten



2. Das Forschungsprojekt ReComMeND



Ziele von ReComMeND

1. Prognosemethodik unter Einsatz mathematischer Modelle optimieren
2. Treibereinflüsse besser verstehen
3. Controlling in Richtung des Predictive Analytics weiterentwickeln
4. Verallgemeinerbare Erkenntnisse für Nah-Verkehrsdienstleistungen ableiten

Projektpartner:



Lufthansa Systems



**Lufthansa
Industry Solutions**

Laufzeit: 1. 10. 2019 – 31. 12. 2022 (beantragt), Budget ca. 300 T€

<https://www.ifaf-berlin.de/projekte/recommend/>



Treiber

Bereits betrachtete Treiber:

- **Bevölkerungswachstum**
- **Witterung** (Wetter)
- **Tourismus** (Übernachtungsanzahl)
- **Beschäftigung / Arbeitslosigkeit**
- **Benzinpreis**
- **Pendler, Auszubildende, Schüler, Studierende** (auf Jahresbasis)

Weitere mögliche Treiber:

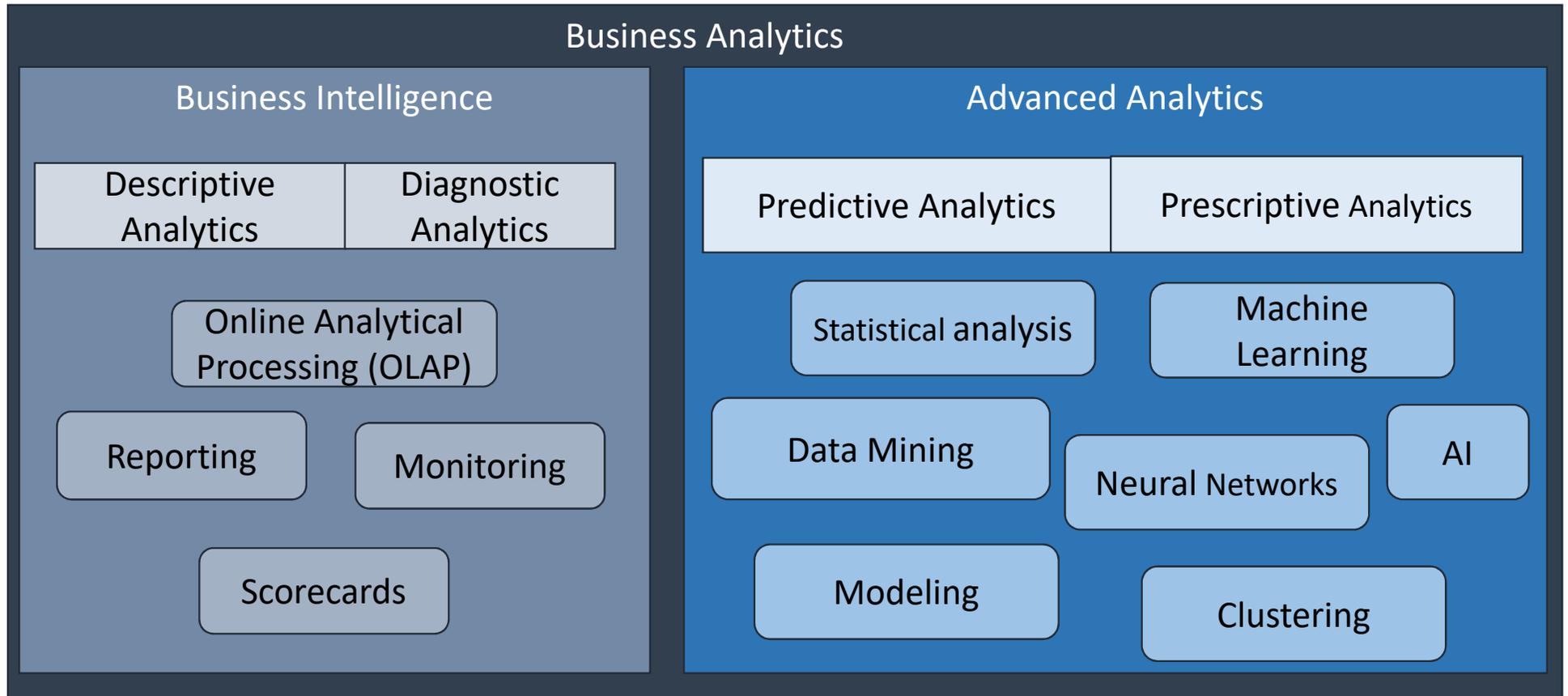
- **Altersstruktur der Bevölkerung**
- **Kalendereffekt**
- **Auswertung von App-Daten**
- **Preisentwicklungen der Fahrscheine**
- **Alternative Mobilitätsangebote**
- **Angebotsorientierte Treiber (Fahrtenangebot)**
- **Kundenorientierte Treiber (Marktforschung)**

Einflüsse durch Schocks

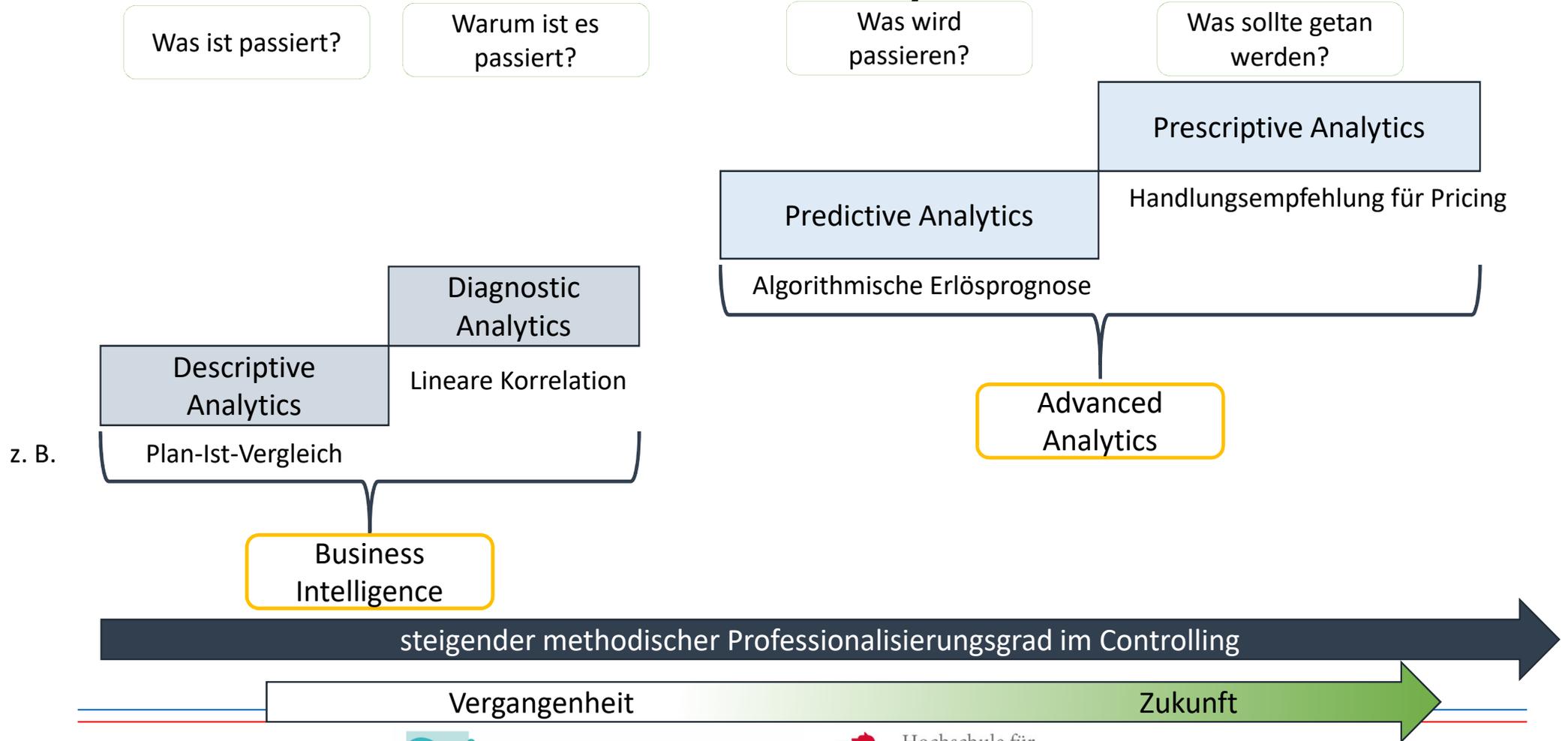
3. Controlling-Kontext



Überblick



Von der BI zu Advanced Analytics



Definitionen

- **Advanced Analytics** umschreibt als **Oberbegriff** Verfahren der Datenanalyse, die über herkömmliche, vorrangig vergangenheitsorientierte deskriptive Methoden der Business Intelligence hinausgehen. Controller/innen müssen ihre statistischen Kompetenzen auffrischen / erweitern, um mit Advanced Analytics-Methoden Handlungsempfehlungen ableiten zu können. Je mehr Datenmanagement vom Controlling selbst (anstelle einer Fachabteilung) verantwortet wird / werden muss, desto größer sind auch die notwendigen Fähigkeiten in Data Science.
- **Predictive Analytics** meint Prognosemodelle, bei denen aus Lernprozessen der Vergangenheit (Künstliche Intelligenz, Maschinelles Lernen) Aussagen über die Zukunft getroffen werden.
- **Prescriptive Analytics** berührt bereits den normativen und interpretatorischen Bereich und leitet aus Datenanalysen Aussagen ab, welche rationale Entscheidungen durch das Unternehmen getroffen werden sollten. Somit berührt Prescriptive Analytics eine Kernkompetenz von Controller/innen und steht potenziell im Wettbewerb um die beste rationale Interpretation.

Quelle: Eigene Synopse



Anwendungsbeispiele ÖPNV

Begriff	Beispiel ÖPNV
Business Intelligence	Monatsbericht (Ist-Werte) zu Erträgen, Aufwendungen, Beförderungsfälle, Personenkilometer, Mengen Ticketverkäufe; Plan-Ist-Abweichungen, Trends; einfache Diagnosen bspw. mit linearer Korrelationsanalyse zwischen zwei Werten (Hotelübernachtungen zu Verkaufsmenge Touristenticket)
Advanced Analytics <i>als Oberbegriff auch inklusive:</i>	Komplexere Korrelationsanalysen zwischen Treibern (Bevölkerung, Wetter, Pendler etc.) zu abhängigen Variablen (Anzahl Ticketverkäufe, Erträge); Regressionsanalysen; Interdependenzen (wie stark wirken sich A, B, C, netto auf X, Y, Z aus? Wie stark beeinflussen sich die Größen untereinander?); algorithmische Verfahren
Predictive Analytics	Kurz- bis langfristige Vorhersagen zu Beförderungsfällen, Ticketverkäufen, Ertragsentwicklung; Simulationen und Szenario-Analysen; Nutzung von verschiedenen Prognosemodellen, bspw. SARIMA, Holt-Winters-Verfahren, exponentielles Glätten, Maschinelles Lernen
Prescriptive Analytics	Algorithmus gibt Empfehlungen ab, wie das Streckennetz ausgebaut werden sollte, wie die Fahrten disponiert werden sollten, wie die Tarife gestaltet werden sollten, wie die Produktkategorien entwickelt werden sollten

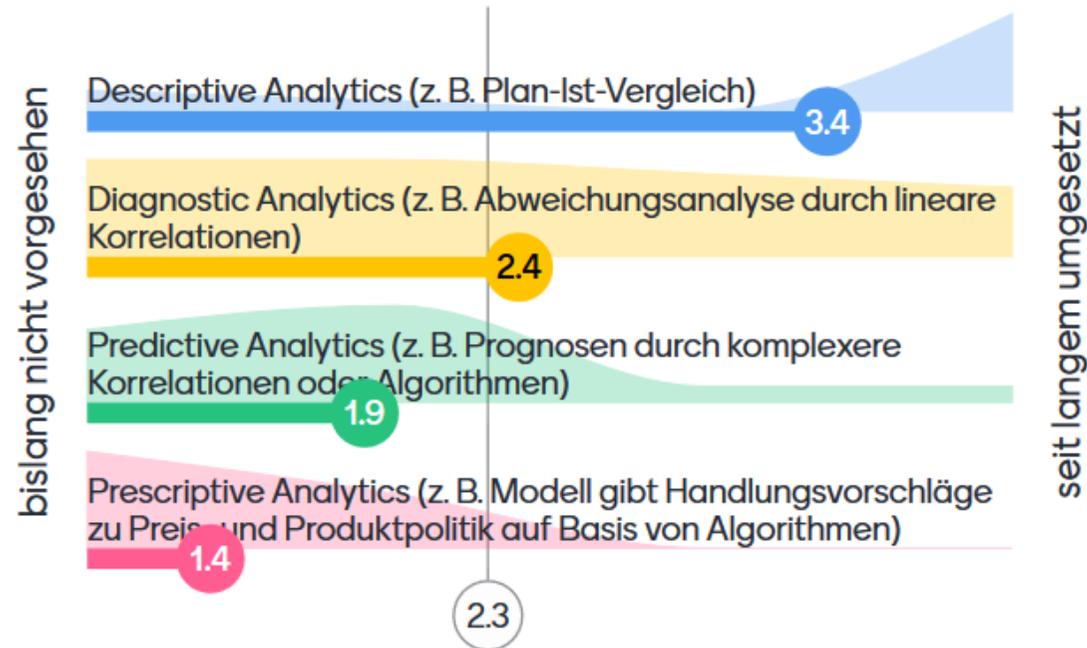
Mentimeter-Umfrage unter Teilnehmer/innen

- QR-Code mit Smartphone scannen und abstimmen:



- Oder auf Link im Zoom-Chat klicken und mit dem Computer abstimmen

Bitte bewerten Sie für das Unternehmen, in dem Sie tätig sind: Wir praktizieren...



Stand: 09.11.2020



Vergleich mit Studie zu Advanced Analytics aus 2018

Welche Analytics-Verfahren finden in Ihrem Unternehmen bereits Anwendung?

66,3% Descriptive Analytics (Was ist passiert?)

60,9% Diagnostische Analytics (Warum ist etwas passiert?)

56,2% Predictive Analytics (Was wird passieren?)

27,2% Prescriptive Analytics (Was muss getan werden?)

Quelle: IDG Studie Predictive Analytics 2018

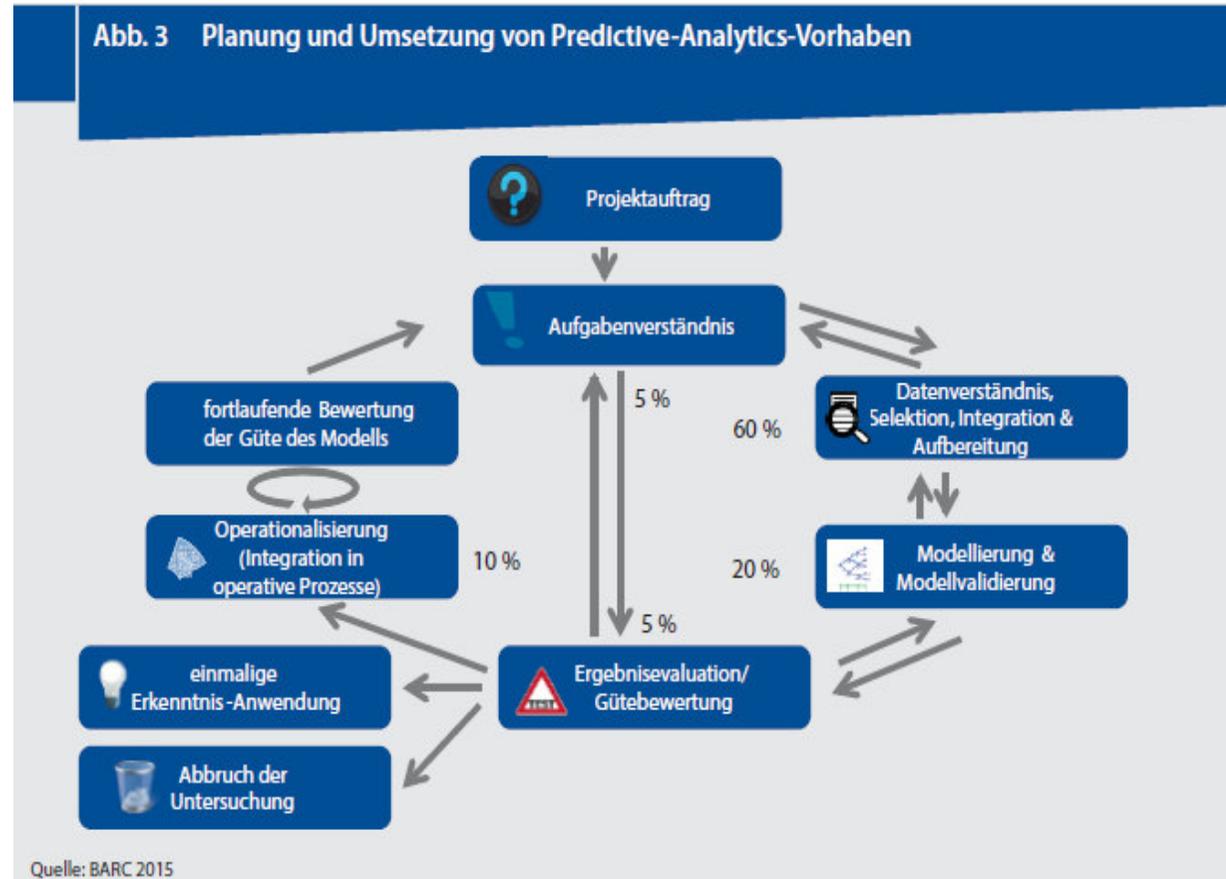
Typische Projektstruktur für Implementierung von Predictive Analytics

Mischung aus:

- KDD-Methodik (Knowledge Discovery in Databases)
- CRISP DM (Cross Industry Standard Process for Data Mining) und
- Erfahrungswissen

Quelle: Iffert, Lars: Predictive Analytics richtig einsetzen. In: Controlling & Management Review, Sonderheft 1 / 2016, S. 20

Abb. 3 Planung und Umsetzung von Predictive-Analytics-Vorhaben



Quelle: BARC 2015

4. Erste Ergebnisse



Datenverständnis und Aufarbeitung

- Monatliche Ertrags-/Erlös- und Verkaufszahlen
- Zeitreihen von 2005 bis 2019
- Über 300 Produkte in dieser Zeit (In 2019 sind es 209)
- 3 Unternehmen (BVG, S-Bahn, DB Regio)
- Tabelle exemplarisch für das Jahr 2019

Produktgruppe	Anzahl Produkte	Anteil Umsatz (in %)
Abonnements	54	43,13
Einzelfahrausweise	21	21,6
Zeitkarten Monat	12	12,17
Tageskarten	15	7,23
Semesterticket	1	4,86
Berlin-Ticket S	1	4,68
Touristenprodukte	95	1,81
Zeitkarten 7-Tage	3	1,57
Zeitkarten Jahr	4	1,06
Kombiniertes Produkt	1	0,3
sonstige	2	1,58
Summe	209	100

Herausforderungen bei der Datenaufarbeitung

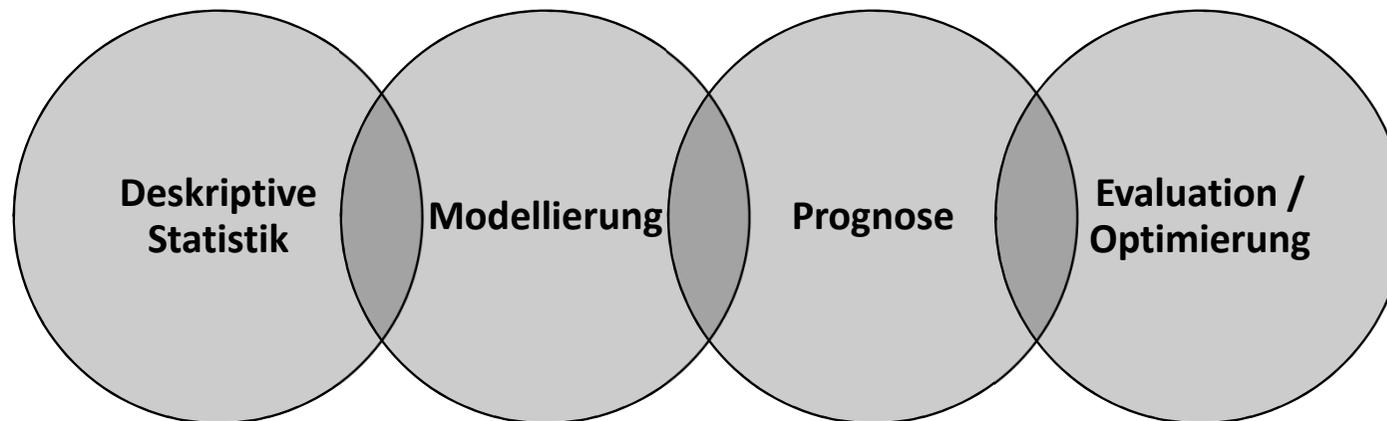
- Zeitpunkt der Ticketnutzung weitestgehend unbekannt
- Erträge vs. Einnahmen:
Nicht einheitlich. Kann sich zwischen den Unternehmen und auch zwischen Produktgruppen unterscheiden
- Fehlende Werte (Bsp.: 3 Monate Lücke bei der DB Regio)
- Abrechnungsmodalitäten (Bsp.: Vertriebspartner Touristenprodukte)
- Semestertickets (Keine Stückzahlen und unregelmäßige Überweisung der Universitäten)
- Bei Nutzergruppen keine eindeutige Zuordnung möglich

Korrelation zwischen Treibern und Produktgruppen

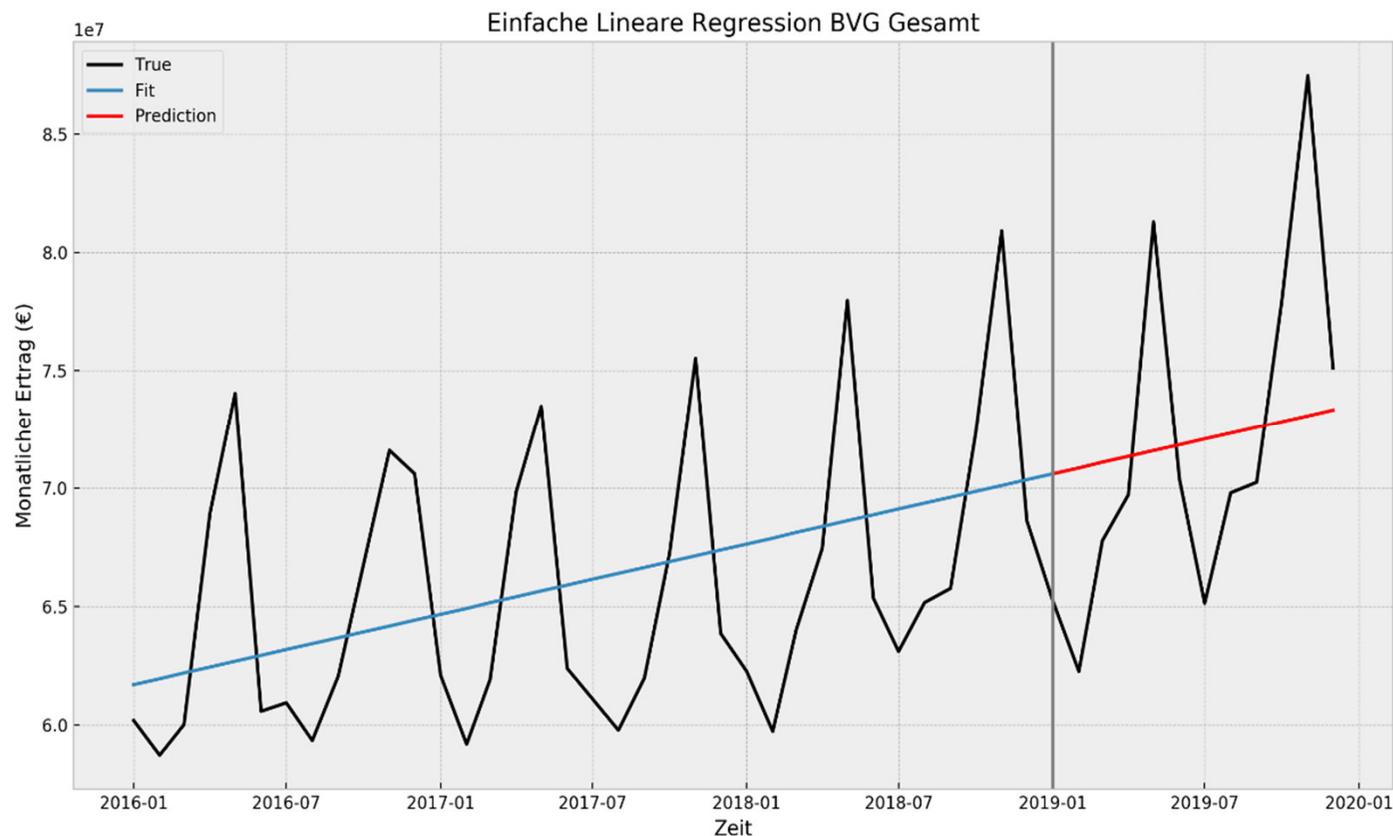
	BVG Gesamt	Abonnements	Einzelfahr- scheine	Zeitkarten Monat	Tageskarten
Arbeitslose	-0,889	-0,901	-0,787	0,285	-0,734
Superbenzin	0,244	0,275	0,154	-0,510 (?)	0,260
Diesel	0,134	0,152	0,047	-0,459 (?)	0,084
Ankünfte	0,839	0,843 (?)	0,730	-0,374	0,921
Übernachtungen	0,797	0,824 (?)	0,706	-0,403	0,950
Bevölkerungs- stand	0,798	0,931	0,676	0,084	0,580

Trainieren und Testen der Modelle

- Autoregressive Modelle werden von 2005-2018 trainiert
- Modelle mit exogenen Variablen (Treibern) werden von 2012-2018 trainiert
- 2019 wird mit den Modellen prognostiziert und mit den realen Werten verglichen

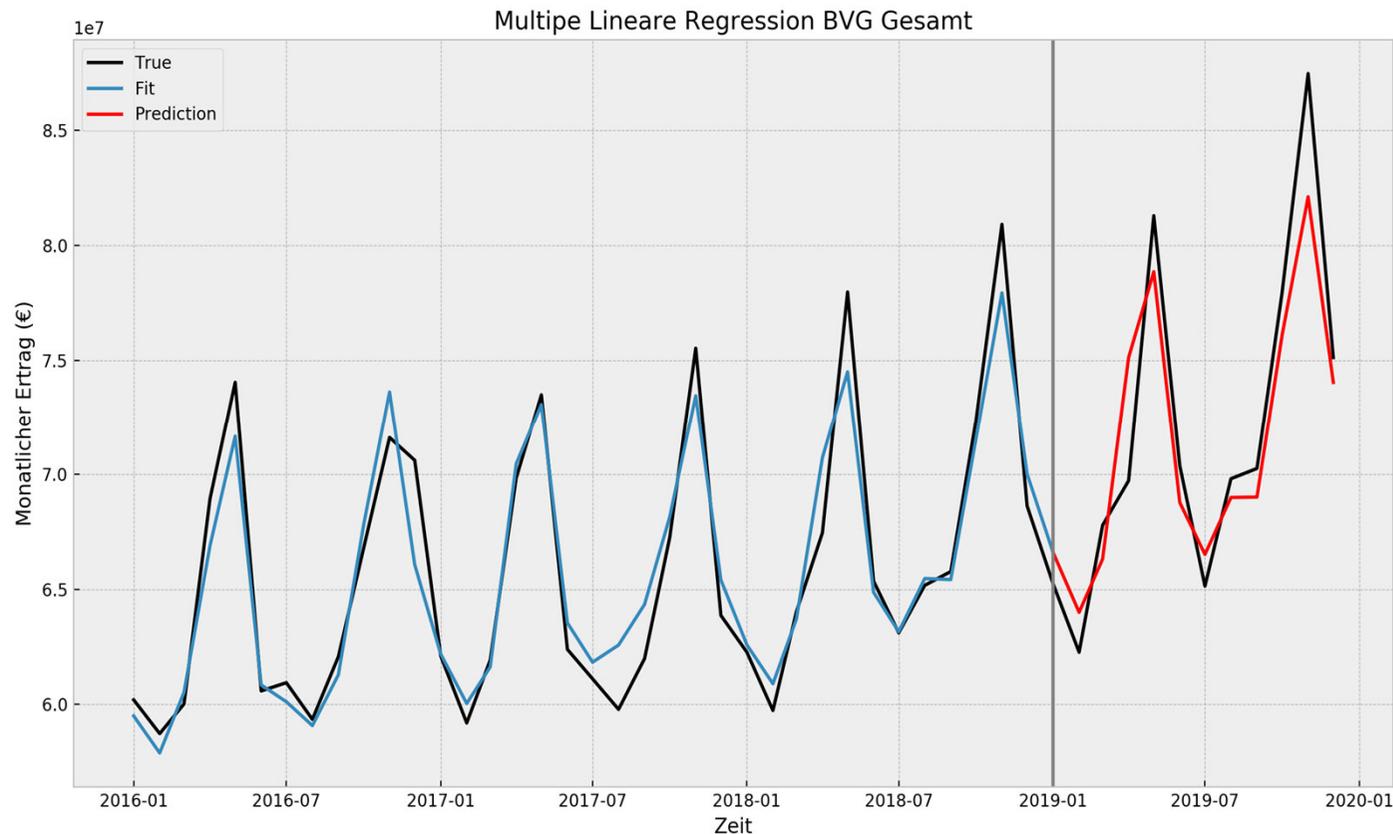


Einfache Lineare Regression der monatlichen Fahrgelderträge der BVG



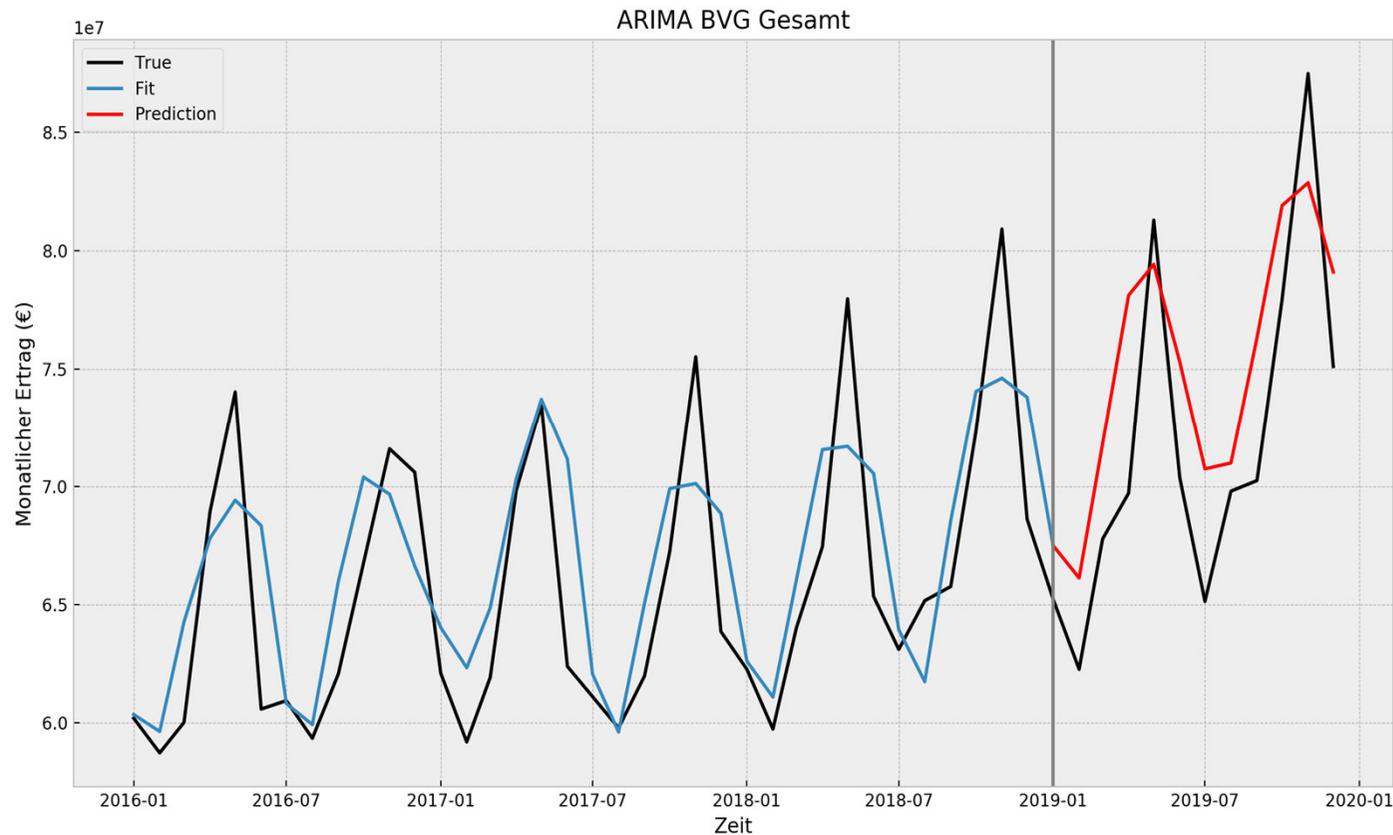
- Mittlerer absoluter prozentualer Fehler
- $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_t - P_t}{A_t} \right|$
Mit A_t aktueller Wert und P_t Prognosewert
- MAPE: 7,19 %

Multivariate Lineare Regression der monatlichen Fahrgelderträge der BVG



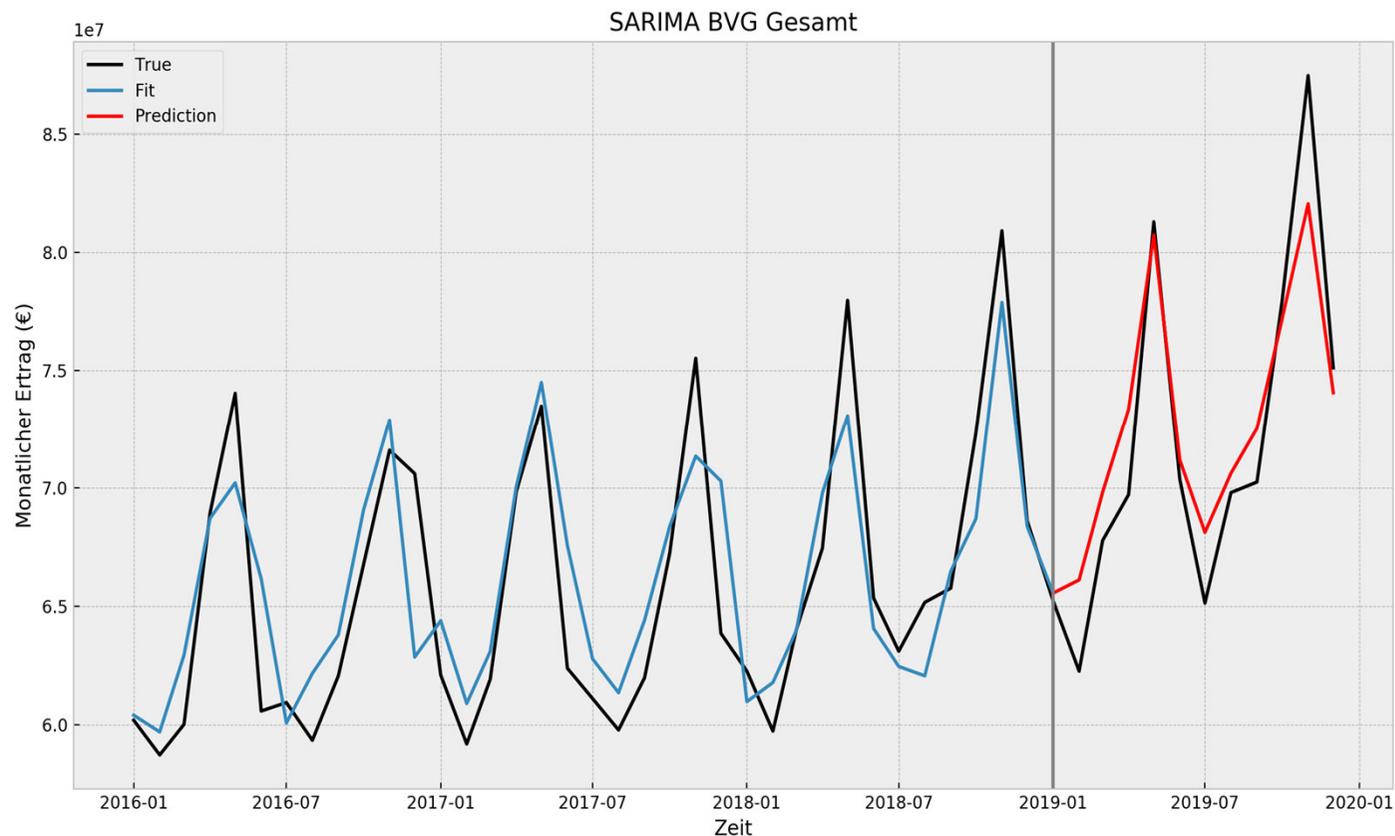
- Übernachtungen
Bevölkerung,
Auszubildende,
Studierende, Schüler,
Arbeitslose,
Dieselpreis,
Monate
- MAPE: 2,92 %

ARIMA-Modell der monatlichen Fahrgelderträge der BVG



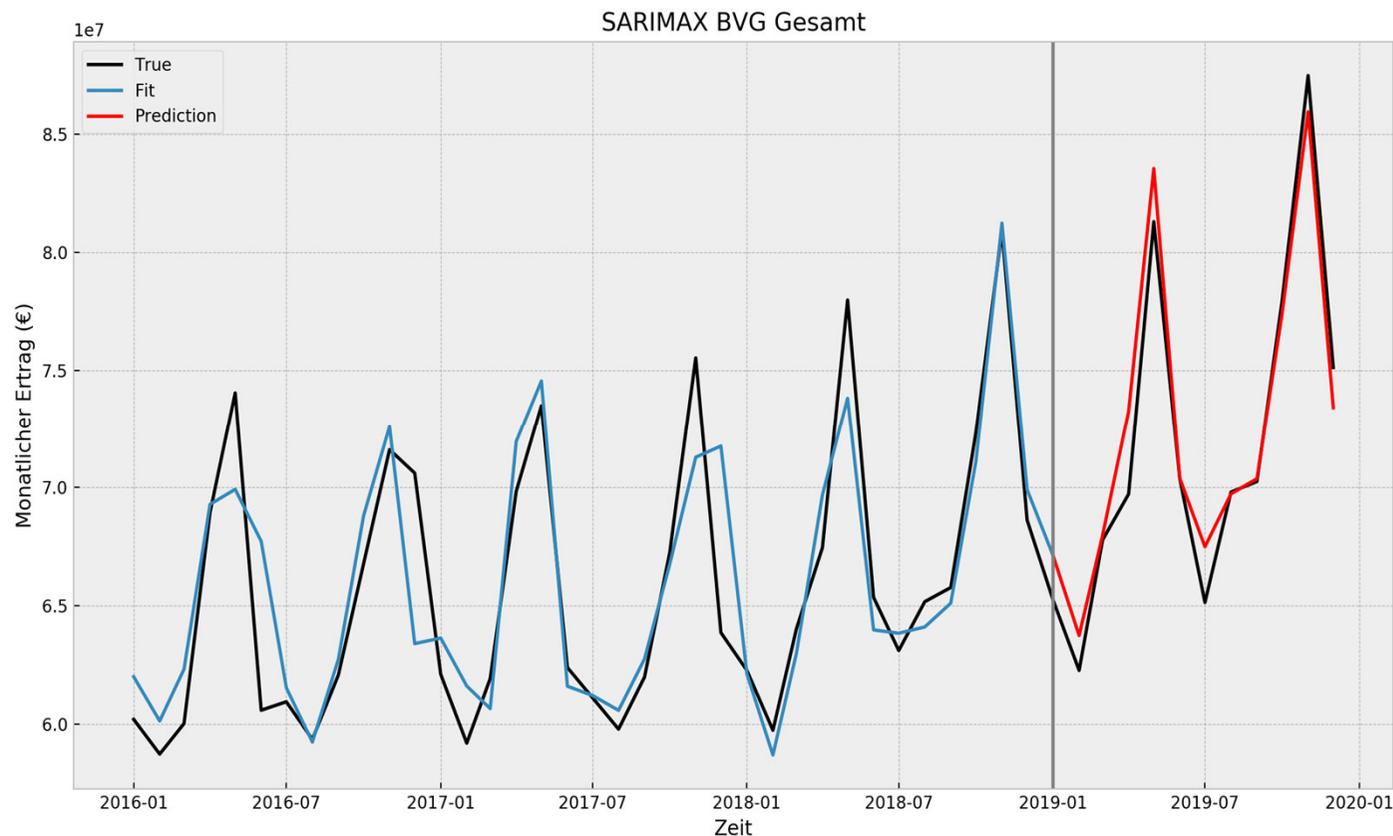
- AR (Autoregressiv)
- I (Integrated)
- MA (Moving Average)
- ARIMA(p,d,q)
- MAPE: 5,96 %

SARIMA-Modell der monatlichen Fahrgelderträge der BVG



- Saisonales ARIMA
- SARIMA
(p,d,q)(P,D,Q)
- MAPE: 2,86 %

SARIMAX-Modell der monatlichen Fahrgelderträge der BVG



- Saisonales ARIMA mit exogenen Variablen
- Übernachtungen, Bevölkerung, Arbeitslose, Dieselpreis
- MAPE: 2,03 %

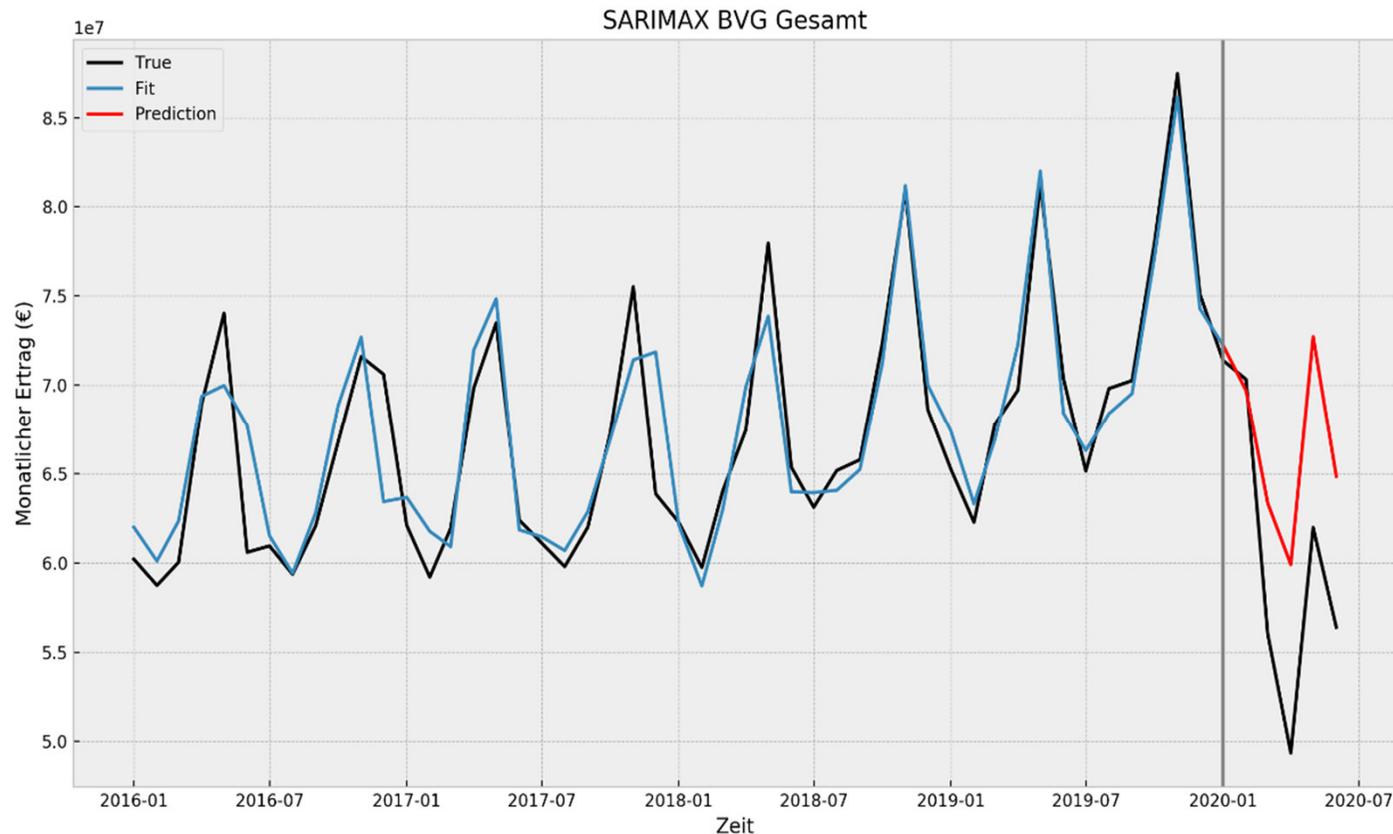
Methodenvergleich mittels MAPE (in %): Prognose zu IST-Werten des Jahres 2019

Methodenvergleich mittels MAPE (in %): Prognose zu IST-Werten des Jahres 2019	BVG Gesamt	Abos	Einzel- fahrtscheine	Monats- karten	Tages- karten
Lineare Regression	7,19	7,28	5,18	18,66	11,12
Multivariate lineare Regression	2,92	8,61	3,07	8,05	6,70
RIDGE Regression	4,70	10,50	2,95	9,81	5,58
LASSO Regression	4,69	10,64	2,66	9,28	4,29
ARIMA	5,96	9,37	5,59	14,58	19,58
SARIMA	2,86	6,67	2,09	9,98	5,21
SARIMAX	2,03	10,78	2,93	11,52	5,11
Holt-Winters	2,62	6,4	3,09	7,09	4,82

Methodenvergleich mittels MAPE (in %): Prognose zu IST-Werten 1. Halbjahres 2020

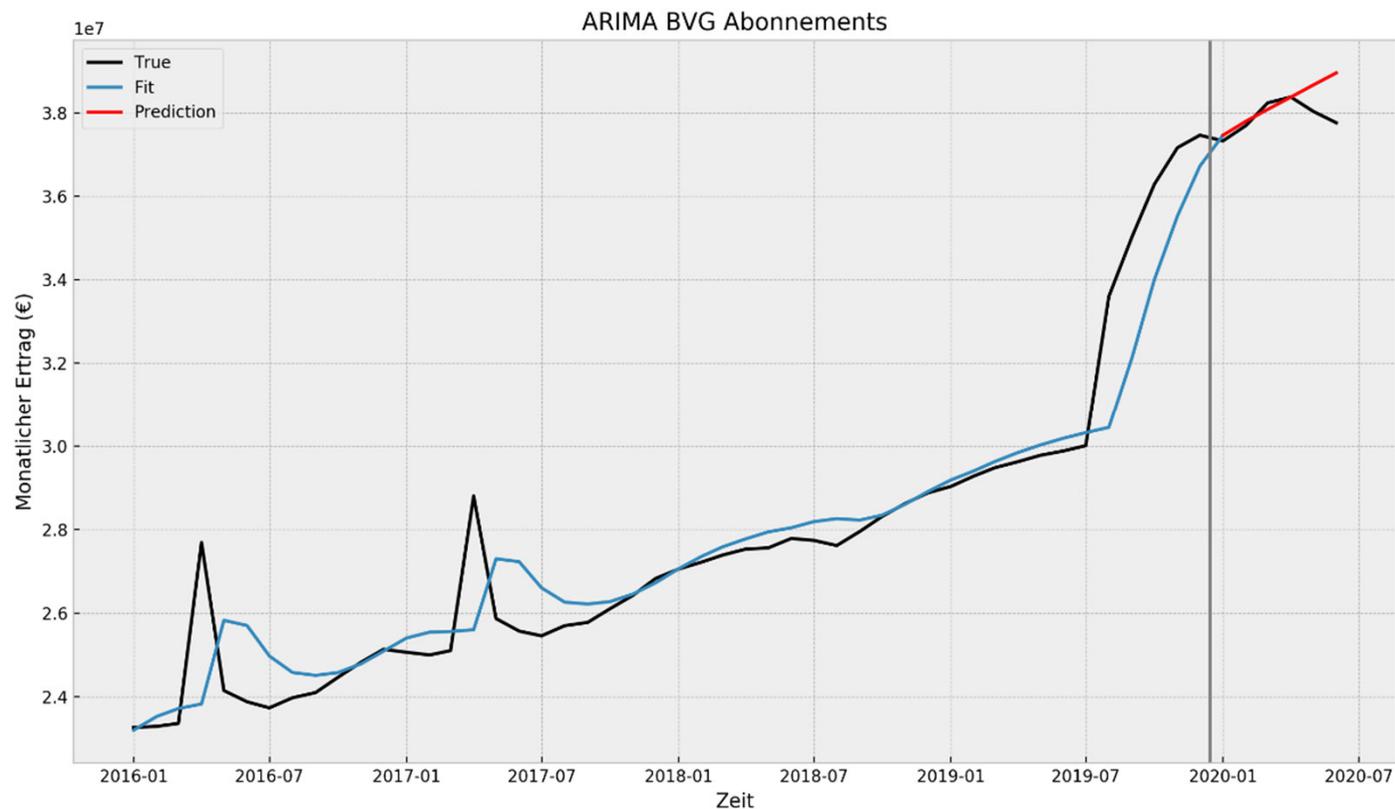
Methodenvergleich mittels MAPE (in %): Prognose zu IST-Werten 1. Halbjahres 2020	BVG Gesamt	Abos	Einzel- fahrtscheine	Monats- karten	Tages- karten
Lineare Regression	21,51	18,12	201,49	128,14	695,25
Multivariate lineare Regression	16,72	18,25	117,75	72,79	130,47
RIDGE Regression	12,7	17,49	154,92	76,1	181,59
LASSO Regression	15,91	9,27	153,84	104,6	121,6
ARIMA	33,73	0,97	193,99	131,39	553,69
SARIMA	35,28	9,12	149,59	111,39	520,65
SARIMAX	11,47	3,89	148,57	154,25	38,06
Holt-Winters	27,46	1,59	186,97	102,62	677,76

SARIMAX-Modell der monatlichen Fahrgelderträge der BVG (Corona)



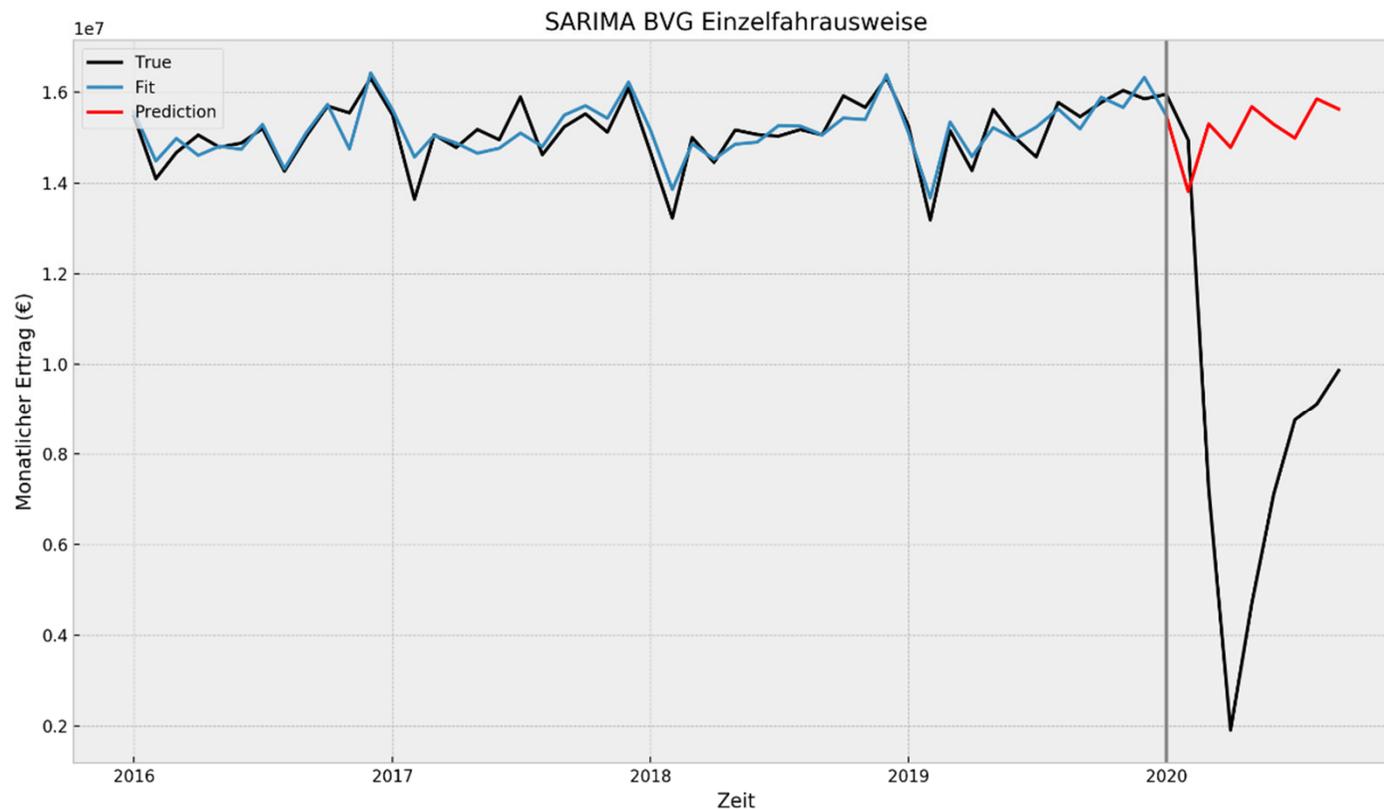
- Prognose wird durch exogene Variablen nach unten korrigiert (Bsp.: Übernachtungen)
- MAPE: 11,47 %

ARIMA-Modell der monatlichen Abonnementserträge der BVG (Corona)



- Keine Exogenen Variablen
- Die Abonnements stagnieren während Corona => keine neuen Kunden, kaum Kündigungen
- ARIMA performt daher außergewöhnlich gut
- MAPE: 0,97 %

SARIMA-Modell der monatlichen Einzelfahrscheinträge der BVG (Corona)



- Keine exogenen Variablen
- Bestes Modell in 2019
- Erträge brechen ein
- Autoregressives Modell stößt an seine Grenzen
- MAPE: 149,59 %

Zwischenfazit zu Prognosen des Corona-Schock

- Autoregressive Modelle liefern bei extremen Schocks bei einer 12-monatigen Prognose keine guten Ergebnisse, bei rollierender Prognose ist eine Adaption möglich.
- Modelle mit exogenen Variablen liefern bessere Ergebnisse.
Problem: Treiberdaten sind am Anfang des Jahres natürlich noch nicht verfügbar -> Einbruch der Übernachtungszahlen war ebenso nicht abzusehen.

Methoden

Univariat	Multivariat
Einfache lineare Regression	Multiple lineare Regression
ARIMA	SARIMAX
SARIMA	Shrinkage Methoden (Ridge, LASSO)
Exponentielles Glätten / Holt-Winters	Regression Trees*
...	Random Forests*
	Neuronale Netze (RNN, LSTM)*

*Ausstehend

5. Ausblick und Einladung



Ausblick “Prepare for your future”

Entwicklung des Rollenbilds und der Aufgaben von Controller/innen

Vom...	Zum...
IT-Anwender	IT- und BI-(Mit)Gestalter; Data Scientist mindestens auf Anwenderniveau
Daten-Lieferanten	Datenverfügbarkeits-Gewährleister (Self Controlling durch Führungskräfte), Daten-Verwalter, Datenqualitätsmanager
Erklärer der Vergangenheit	Prognosegeber der Zukunft
Excel-Anwender	Nutzung von modernen Advanced Analytics-Methoden inkl. Algorithmen
IT-Nutzer, IT-Servicebezieher	(Mit-)Gestaltung bis fachliche Leitung von IT-Projekten im Controlling-Umfeld
Verwalter einer noch intuitiv fassbaren Datenmenge	Produzent von methodensicheren Big Data-Analysen mit Einsatz statistischer Methodik

Einladung zum Praxisaustausch

Als konkrete Fragestellungen für die nächsten Workshops wurden herausgearbeitet:

- In welchen Verkehrsbetrieben und mit welchen Erfahrungen werden bereits mathematische/statistische Prognosemodelle bzw. -methoden für die Erlösplanung eingesetzt?
- Welche Treiber zur Erlösplanung werden von Verkehrsbetrieben a) als relevant angesehen und b) in ihren Controlling-Prozessen bereits berücksichtigt?
- Welche Informationen aus dem Marketing könnten darüber hinaus zur Verfügung stehen, um Erlösprognosen zu verbessern? Welche davon werden im Controlling genutzt?

Haben wir Ihr Interesse am Austausch geweckt? Sie verfügen über einschlägige Erfahrungen?

Dann laden wir Sie ein, sich per Email an den Arbeitsgruppensprecher Kyanusch Kay des ICV zu wenden:

ICV-Bbg.Kay@email.de

Wir freuen uns auf den Austausch mit Ihnen!

Kontakte

Projekt-Exposé mit allen Beteiligten: <https://www.ifaf-berlin.de/projekte/recommend/>

Prof. Dr. Robert Knappe

Professur für Betriebswirtschaftslehre der öffentlichen Verwaltung

Jonas Krembsler

Wissenschaftlicher Mitarbeiter im Projekt ReComMeND

Hochschule für Wirtschaft und Recht Berlin

Alt Friedrichsfelde 60

10315 Berlin

Tel. (+49) (0)30 - 30877 - 2610

Email: robert.knappe@hwr-berlin.de bzw. jonas.krembsler@hwr-berlin.de

