

WIRTSCHAFTSINFORMATIK 2

DATENBANKEN + EXCEL, DATENANALYSE

BERND BLÜMEL, CHRISTIAN BOCKERMANN,
CHRISTIAN METZGER

HOCHSCHULE BOCHUM

SOMMERSEMESTER 2023

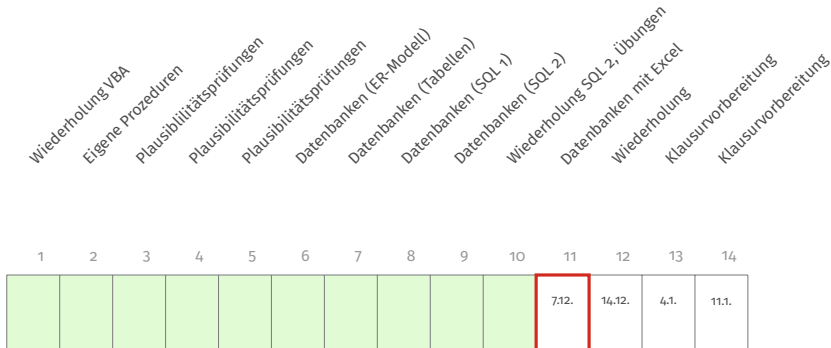
Bisherige Vorlesung

Wiederholung VBA	Eigene Prozeduren	Plausibilitätsprüfungen	Plausibilitätsprüfungen	Plausibilitätsprüfungen	Datenbanken (ER-Modell)	Datenbanken (Tabellen)	Datenbanken (SQL 1)	Datenbanken (SQL 2)	Wiederholung SQL 2, Übungen	Datenbanken mit Excel	Wiederholung	Klausurvorbereitung	Klausurvorbereitung
1	2	3	4	5	6	7	8	9	10	11	12	13	14
										7.12.	14.12.	4.1.	11.1.

Bisherige Vorlesung

Wiederholung VBA	Eigene Prozeduren	Plausibilitätsprüfungen	Plausibilitätsprüfungen	Plausibilitätsprüfungen	Datenbanken (ER-Modell)	Datenbanken (Tabellen)	Datenbanken (SQL 1)	Datenbanken (SQL 2)	Wiederholung SQL 2, Übungen	Datenbanken mit Excel	Wiederholung	Klausurvorbereitung	Klausurvorbereitung
1	2	3	4	5	6	7	8	9	10	11	12	13	14
										7.12.	14.12.	4.1.	11.1.

Bisherige Vorlesung

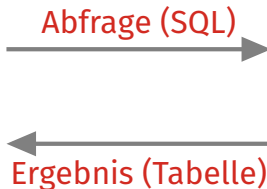
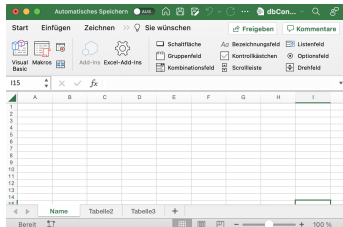


<https://datascience.hs-bochum.de>

→ Vorlesung → Wirtschaftsinformatik 2

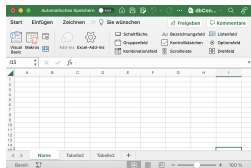
Datenbank-Zugriff mit Excel

- SQL Abfragen von VBA aus starten
- Ergebnis kann dann in Excel-Tabelle geschrieben werden



Datenbank

ODBC Treiber



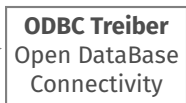
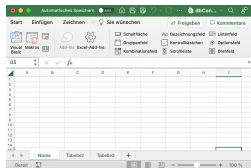
ODBC Treiber
Open DataBase
Connectivity



Datenbank

- ODBC ist Vermittlungssoftware zwischen Excel und DB
- Treiber muss im Betriebssystem installiert werden
- Treiber spezifisch für jeweilige Datenbank

ODBC Treiber



Datenbank

- ODBC ist Vermittlungssoftware zwischen Excel und DB
- Treiber muss im Betriebssystem installiert werden
- Treiber spezifisch für jeweilige Datenbank
- MySQL Datenbank -> ODBC für MySQL, usw.

SQL Aufruf in VBA

Verbindungsstring erzeugen:

```
Dim conn As ADODB.Connection
Dim url As String
Dim rec As ADODB.Recordset
Dim SQL As String
Dim i As Long
i = 2
```

```
url = "Driver=MYSQL ODBC 8.0 Unicode Driver"
url = url & ";Server=pav050.hs-bochum.de"
url = url & ";Database=KundenDB"
url = url & ";UID=wiInf"
url = url & ";PWD=wiInf"
```


SQL Aufruf in VBA

```
Set conn = New ADODB.Connection
conn.ConnectionString = url
conn.Open
SQL = "Select Name, PLZ FROM Kunde"
Set rec = New ADODB.Recordset

rec.Open SQL, conn
Do While Not rec.EOF
    Sheets(1).Cells(i, 1) = rec!Name
    Sheets(1).Cells(i, 2) = rec!PLZ
    rec.MoveNext
    i = i + 1
Loop
```

Problem: Kein frei verfügbarer ODBC-Treiber für Mac OS

- ODBC-Zugriff funktioniert unter Windows
- ODBC-Treiber für Mac OS kostenpflichtig
- Keine Klausuraufgabe zu Datenbankzugriff mit Excel!

Datenanalyse

Vom Geschäftsprozess zur Datenbank

Geschäftsprozess



Warenkorb			
Kunden-Nr.	20070		
Bestellzeit	04.11.2016 08:07		
Artikel	Bestandshöhe	Preis	
100	1	10,00 €	
200	1	30,00 €	
300	1	10,00 €	
In 20% Quantitäten			
Versandkosten		20,00 €	
Gesamtbetrag		70,00 €	

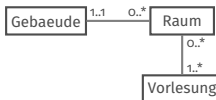
“Student:in besucht Vorlesung.”

“Gebäude haben
mindestens 1 Raum.”

Interviews mit
Fachabteilungen

Textuelle Beschreibungen,
Use-Case Diagramme

Entity-Relationship Modell



Entitätstypen

Relationen

ER Diagramm
in der Sprache UML

Datenbank

Bestellung

BestellNr	Datum	Kundennr
1	2016-05-01	1
2	2016-05-04	2
3	2016-05-09	3
4	2016-05-17	1

Artikel

ArtikelNr	Bezeichnung	Preis
1	Lasermouse	17,99
2	Tastatur	27,99
3	Flexplatte	59,99
4	USB-Stick	9,99
5	WebCam	19,99

Kunde

Kundennr	Name	Vorname	PLZ	Ort
1	Maier	Mia	44801	Bochum
2	Schmidt	Jonas	44801	Bochum
3	Müller	Emma	44800	Bochum
4	Weber	Lukas	45127	Essen

SQL Schema Definition

Tabellen, Relationstabellen,
Constraints/Regeln

Beispiel: Analyse von Studierenden-Zahlen



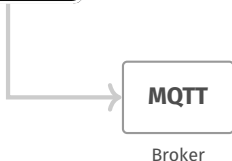
252,	6.12.2022,	9:45,	AW-2-33,	+1
253,	6.12.2022,	9:47,	AW-2-33,	+1
254,	6.12.2022,	9:48,	AW-2-33,	-1
255,	6.12.2022,	9:51,	AW-2-33,	+1
256,	6.12.2022,	9:51,	AW-2-33,	+1
258,	6.12.2022,	9:51,	AW-2-33,	+1
259,	6.12.2022,	9:52,	AW-2-33,	+1

Annahme: Genügend Räume sind mit einem Zähler versehen.

Beispiel: Analyse von Studierenden-Zahlen



```
252, 6.12.2022, 9:45, AW-2-33, +1
253, 6.12.2022, 9:47, AW-2-33, +1
254, 6.12.2022, 9:48, AW-2-33, -1
255, 6.12.2022, 9:51, AW-2-33, +1
256, 6.12.2022, 9:51, AW-2-33, +1
258, 6.12.2022, 9:51, AW-2-33, +1
259, 6.12.2022, 9:52, AW-2-33, +1
```

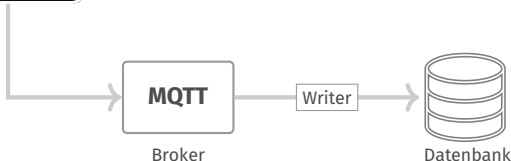


Annahme: Genügend Räume sind mit einem Zähler versehen.

Beispiel: Analyse von Studierenden-Zahlen

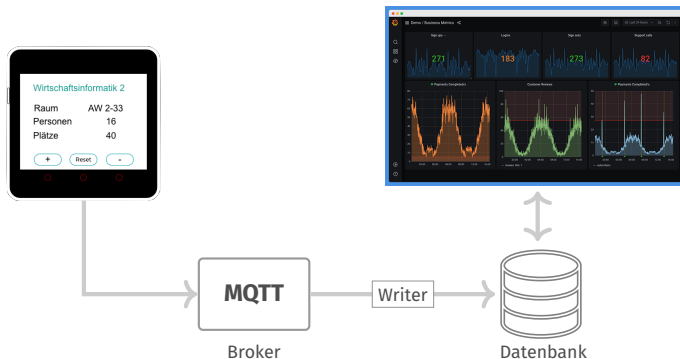


252,	6.12.2022,	9:45,	AW-2-33,	+1
253,	6.12.2022,	9:47,	AW-2-33,	+1
254,	6.12.2022,	9:48,	AW-2-33,	-1
255,	6.12.2022,	9:51,	AW-2-33,	+1
256,	6.12.2022,	9:51,	AW-2-33,	+1
258,	6.12.2022,	9:51,	AW-2-33,	+1
259,	6.12.2022,	9:52,	AW-2-33,	+1



Annahme: Genügend Räume sind mit einem Zähler versehen.

Beispiel: Analyse von Studierenden-Zahlen

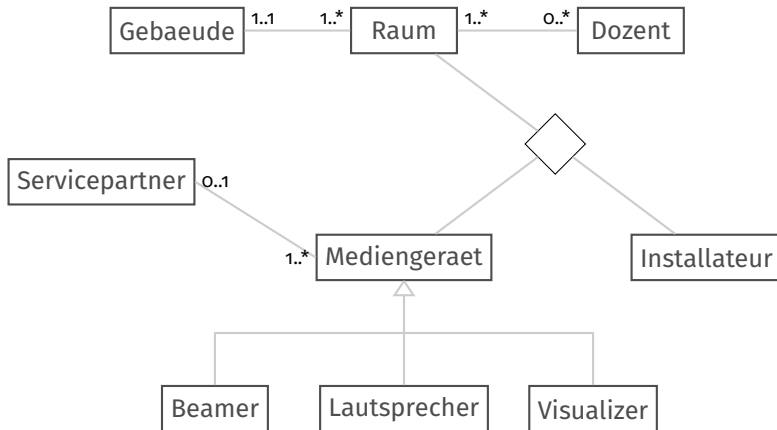


Beispiel: Analyse von Studierenden-Zahlen



252,	6.12.2022,	9:45,	AW-2-33,	+1
253,	6.12.2022,	9:47,	AW-2-33,	+1
254,	6.12.2022,	9:48,	AW-2-33,	-1
255,	6.12.2022,	9:51,	AW-2-33,	+1
256,	6.12.2022,	9:51,	AW-2-33,	+1
258,	6.12.2022,	9:51,	AW-2-33,	+1
259,	6.12.2022,	9:52,	AW-2-33,	+1

- Was wären spannende Fragen, die wir mit diesen Daten beantworten könnten?
- Wie speichern wir diese Daten so, dass wir *gut* darauf zugreifen können?





Tabellen für die Basis Entitäten

Tabellenname	Primärschlüssel	weitere Attribute
Gebaeude	GebaeudeNr	Bezeichnung, Strasse, PLZ, Ort
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u>
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u>

Frage: Wie speichern wir die Raumauslastung aus den Sensor-Daten?

Tabellen für die Basis Entitäten

Tabellenname	Primärschlüssel	weitere Attribute
Gebaeude	GebaeudeNr	Bezeichnung, Strasse, PLZ, Ort
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u>
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u>

Frage: Wie speichern wir die Raumauslastung aus den Sensor-Daten?

Tabellenname	Primärschlüssel	weitere Attribute
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u> , Anzahl
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u> , Anzahl

Speichern der Sensordaten



252, 6.12.2022, 9:45, AW-2-33, +1
253, 6.12.2022, 9:47, AW-2-33, +1
254, 6.12.2022, 9:48, AW-2-33, -1

```
UPDATE Raum SET Anzahl = Anzahl + WERT  
WHERE Raum.name = 'AW-2-33'
```

Speichern der Sensordaten



252, 6.12.2022, 9:45, AW-2-33, +1
253, 6.12.2022, 9:47, AW-2-33, +1
254, 6.12.2022, 9:48, AW-2-33, -1

```
UPDATE Raum SET Anzahl = Anzahl + WERT  
WHERE Raum.name = 'AW-2-33'
```

Keine gute Idee?

Tabellen für die Basis Entitäten

Tabellenname	Primärschlüssel	weitere Attribute
Gebaeude	GebaeudeNr	Bezeichnung, Strasse, PLZ, Ort
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u>
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u>

Frage: Wie speichern wir die Raumauslastung aus den Sensor-Daten?

Tabellen für die Basis Entitäten

Tabellenname	Primärschlüssel	weitere Attribute
Gebaeude	GebaeudeNr	Bezeichnung, Strasse, PLZ, Ort
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u>
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u>

Frage: Wie speichern wir die Raumauslastung aus den Sensor-Daten?

Tabellenname	Primärschlüssel	weitere Attribute
ZaehlerMeldung	MeldungNr	Datum, Uhrzeit, <u>RaumNr</u> , Wert

Tabellen für die Basis Entitäten

Tabellenname	Primärschlüssel	weitere Attribute
Gebaeude	GebaeudeNr	Bezeichnung, Strasse, PLZ, Ort
Raum	RaumNr	Name, Groesse, <u>GebaeudeNr</u>
Vorlesung	VorlesungNr	Titel, Tag, Zeitslot, <u>RaumNr</u>

Frage: Wie speichern wir die Raumauslastung aus den Sensor-Daten?

Tabellenname	Primärschlüssel	weitere Attribute
ZaehlerMeldung	MeldungNr	Datum, Uhrzeit, <u>RaumNr</u> , Wert, Tag, Zeitslot

Speichern der Messwerte

Berechne Tag, Zeitslot für eingehende Messwerte, dann:

```
INSERT INTO ZahlerMeldung VALUES  
  (id, datum, uhrzeit, raum_nr, wert, tag,  
                                     zeitslot)
```

Wöchentliche Teilnehmerzahlen für *WInf 2*

```
SELECT YEAR(z.datum), YEARWEEK(z.datum), SUM(wert)
FROM Zaehlermeldung z
LEFT JOIN Vorlesung v
ON z.tag = v.tag
AND z.zeitslot = v.zeitslot
AND z.raum_nr = v.raum_nr
WHERE
v.titel = 'Wirtschaftsinformatik 2'
GROUP BY YEAR(z.datum), YEARWEEK(z.datum)
```

Hinweis: Syntax mit YEARWEEK funktioniert by MySQL

Wie geht's von hier aus weiter?

- Wie ist der durchschnittliche Verlauf von Studierendenzahlen über das Semester?
- Vergleich zu Vorjahreszahlen? Trends?
- Wann sind die Tagen mit den meisten Besuchern?
- Gab es seit Schließung der Mensa einen Einbruch der Zahlen?
- Zusammenhang zw. Besucherzahlen und Bahnproblemen/Stau/Wetter?

Problem: Besucherzahlen eher unspezifisch

- Personengenaue Erfassung? Anonym?
- Welche Alternativen gibt es?

Problem: Besucherzahlen eher unspezifisch

- Personengenaue Erfassung? Anonym?
- Welche Alternativen gibt es?



00:00:5e:00:53:af
MAC Adresse



Access Point
ID: AP_812474
Raum: AW 2-33

Ziel: Zur Verbesserung des Lehrangebots

Wozu würden Sie zustimmen?

- Zustimmung MAC Adresse (anonymisiert) im Raum

Ziel: Zur Verbesserung des Lehrangebots

Wozu würden Sie zustimmen?

- Zustimmung MAC Adresse (anonymisiert) im Raum
- Studiengang?

Ziel: Zur Verbesserung des Lehrangebots

Wozu würden Sie zustimmen?

- Zustimmung MAC Adresse (anonymisiert) im Raum
- Studiengang?
- Fachsemester?

Ziel: Zur Verbesserung des Lehrangebots

Wozu würden Sie zustimmen?

- Zustimmung MAC Adresse (anonymisiert) im Raum
- Studiengang?
- Fachsemester?
- Geburtsjahr?

Ziel: Zur Verbesserung des Lehrangebots

Wozu würden Sie zustimmen?

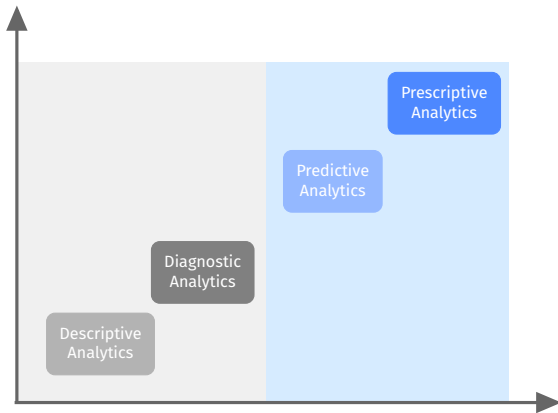
- Zustimmung MAC Adresse (anonymisiert) im Raum
- Studiengang?
- Fachsemester?
- Geburtsjahr?
- Geburtstag?

Ziel: Zur Verbesserung des Lehrangebots

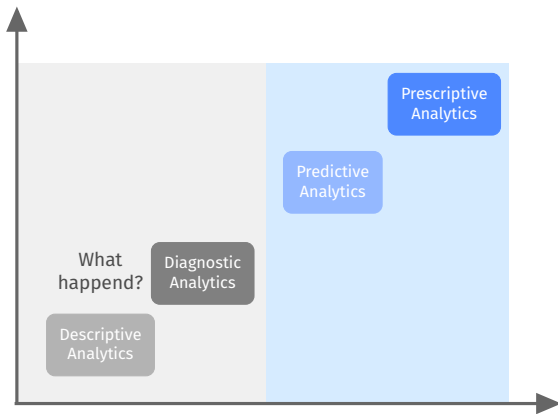
Wozu würden Sie zustimmen?

- Zustimmung MAC Adresse (anonymisiert) im Raum
- Studiengang?
- Fachsemester?
- Geburtsjahr?
- Geburtstag?
- Kontaktdaten?

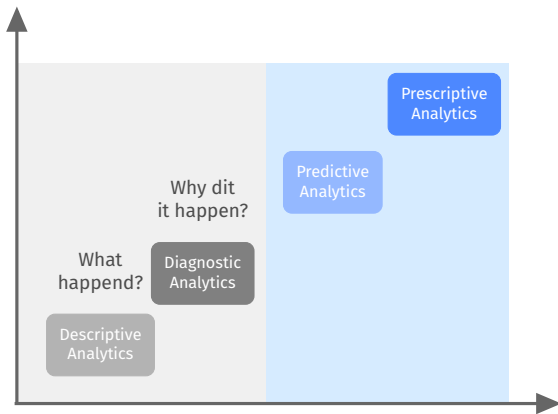
Data Science bzw. *Advanced Analytics* als Weiterentwicklung von *Business Intelligence*



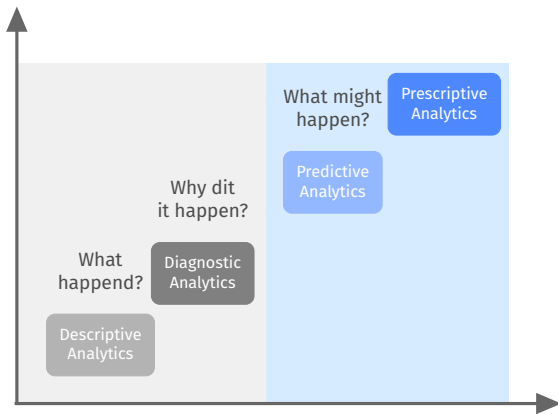
Data Science bzw. *Advanced Analytics* als Weiterentwicklung von *Business Intelligence*



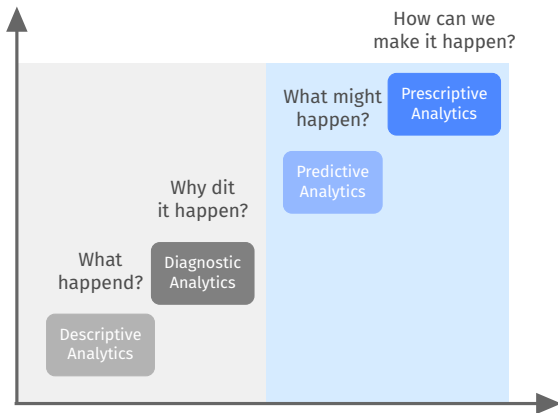
Data Science bzw. *Advanced Analytics* als Weiterentwicklung von *Business Intelligence*

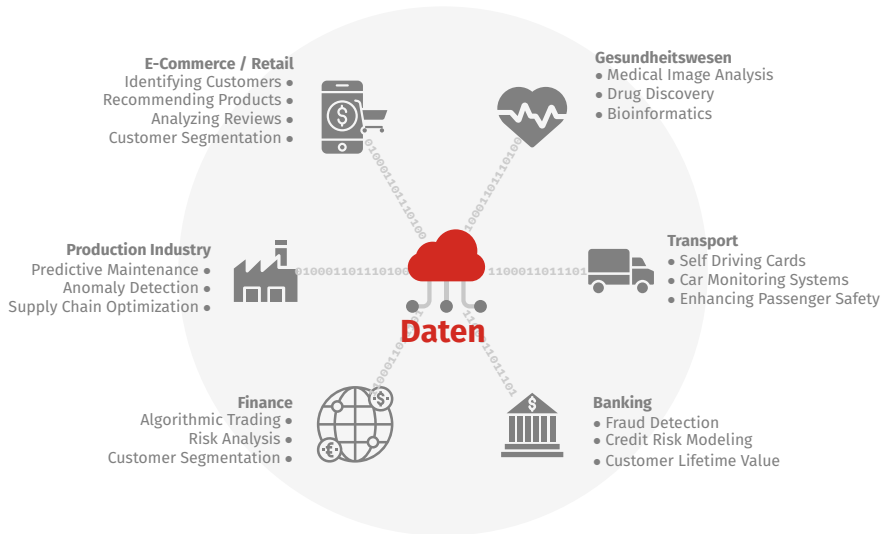


Data Science bzw. *Advanced Analytics* als Weiterentwicklung von *Business Intelligence*

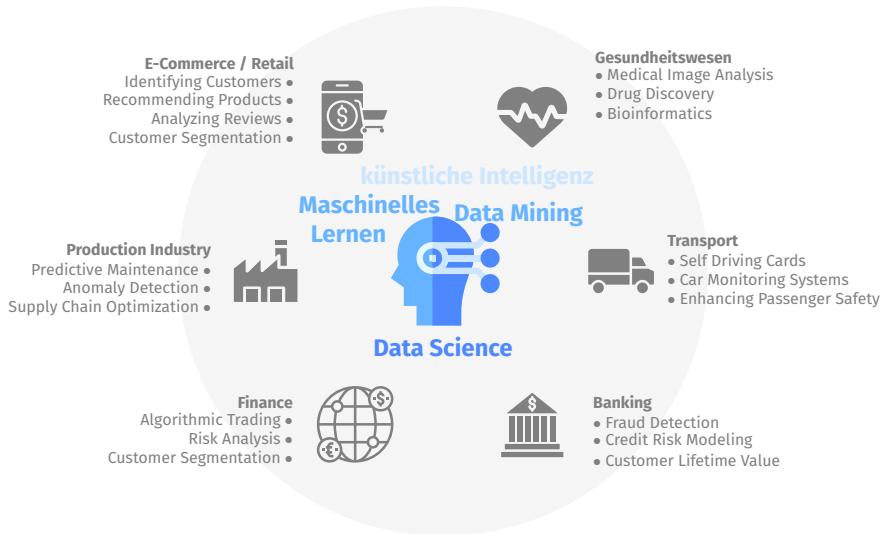


Data Science bzw. *Advanced Analytics* als Weiterentwicklung von *Business Intelligence*





[<https://data-flair.training/blogs/data-science-applications/>]



[<https://data-flair.training/blogs/data-science-applications/>]

Datengetriebene Unternehmenskultur/-philosophie



Wie können Unternehmensdaten zu besseren wirtschaftlichen Entscheidungen genutzt werden?

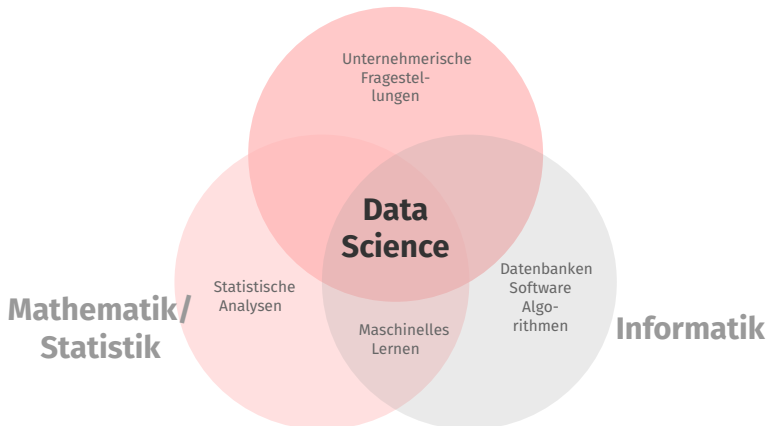


Welche wirtschaftliche Wertschöpfung kann aus Daten erzielt werden?

Data Science ist die Kunst, aus ziemlich viel Datenmaterial wichtige **Erkenntnisse für Entscheider** zu ziehen [...].

[www.online-gruender.com]

Wirtschaft / Anwendung



Was vermittelt der Kurs Data Science?

Ausgehend von unternehmerischen Fragestellungen, werden sämtliche Schritte der Datenanalyse vorgestellt und **selbstständig angewendet**

