

PROZESSE UND UNTERNEHMENSSTRUKTUR

Data Mining

0

Data Mining Einführung

Data Mining

Einführung



Quelle: https://stmaaprodfwsite.blob.core.windows.net/assets/sites/1/2023/03/tesco-shop-AdobeStock_548515188_-william-adobestock.jpg

- Beispiel aus dem Buch "Creating Value with Big Data Analysis"
- Tesco setzte sich mit Einkäufen auseinander, die mit der Club Card getätigt wurden
- Stellten fest, dass Kunden neben Windeln oft auch Bier kaufen

Data Mining

Einführung



Quelle: https://stmaaprodfwsite.blob.core.windows.net/assets/sites/1/2023/03/tesco-shop-AdobeStock_548515188_-william-adobestock.jpg

- Zusätzlich wurde Freitagabends vermehrt Bier und Chips gekauft
- Somit konnte Tesco sein Marketing anpassen und optimieren

1

Data Mining

Was ist Data Mining?

Data Mining beschreibt den Versuch, in Daten Muster zu finden und nutzbar zu machen. Diese Muster oder empirische Zusammenhänge sollen dabei möglichst automatisiert durch Algorithmen erkannt werden. Üblicherweise werden so große Datenmengen analysiert.

Data Mining

Was ist Data Mining?

- Data Mining wird oft mit Big Data gleichgesetzt
- Big Data bezieht sich jedoch primär auf die Verarbeitung umfangreicher Datenbestände und wie die Technik diese ermöglicht
- Allerdings ist Data Mining lediglich ein Werkzeug, um Big Data nutzbar zu machen
- Wird auf riesige Datenbanken angewendet
- Data Mining sucht also in dem Big Data Pool nach verwertbaren Informationen und schafft so Wissen, das sonst ohne Hilfe in großen Datensätzen kaum noch zu erkennen ist

Ziel ist es, Dinge zu entdecken, die neue Erkenntnisse liefern und die dabei helfen, **die eigene Business-Strategie zu verbessern**. Es können also beispielsweise Verbindungen zwischen verschiedenen Produkten auftauchen, die öfter miteinander gekauft werden. Damit kannst du dann Cross-Selling effektiver einsetzen.

2

Data Mining Vorteile im E-Commerce

Neue Fragen oder Probleme aufdecken

Bei Data Mining geht es im Grunde darum, Antworten auf Fragen zu finden, von denen du gar nicht wusstest, dass du diese überhaupt hattest.

A large, solid orange shape that curves from the bottom right towards the center of the slide.

01

Kundenverhalten verstehen

Data Mining ermöglicht die Analyse von Kundenverhalten, Einkaufs-Mustern und Trends, um zukünftige Einkäufe und Präferenzen besser vorhersagen zu können.

02

03

Durch die Segmentierung von Kunden und die Identifikation von Käufertypen können Marketingkampagnen und Angebote individuell angepasst und personalisiert werden.

Personalisiertes Marketing

Data Mining hilft bei der Vorhersage von Verkaufstrends, um den Lagerbestand effizient zu verwalten und Überbestände oder Stock-Outs zu vermeiden.

**Lagerbestand
optimieren**

Produktempfehlungen

Auf Basis von Kaufhistorie und -präferenzen können individuelle Produktempfehlungen generiert werden, was die Möglichkeiten im Cross- und Up-Selling erhöht.

Kundenbindung und -zufriedenheit

Durch das Verständnis der Bedürfnisse und Wünsche der Kunden können Service- und Produktangebote verbessert und die Kundenbindung und -zufriedenheit gesteigert werden.

3

Data Mining Anwendungsbeispiele

Data Mining

Anwendungsbeispiele



Prognosen zur Customer Lifetime Value

Den wahrscheinlichen Ertrag eines Kunden zu kennen, ermöglicht es, die Werbeaufwendungen bei der Kundengewinnung individuell auf bestimmte Zielgruppen zuzuschneiden.



Optimierung von Cross- und Upselling

Data Mining kann erfolgversprechende Muster in Cross- und Up-Selling-Kampagnen finden und so Kunden zur Erhöhung ihres Warenkorbwerts animieren.



Kundensegmentierung

Eine gute Untergliederung der Kundendaten führt zu einem besseren Kundenverständnis und legt somit den Grundstein für effektive Marketingmaßnahmen und eine starke Kundenbindung.

Data Mining

Anwendungsbeispiele



Kampagnenoptimierung

Data Mining hilft bei der Kampagnenoptimierung durch die verbesserte Kundensegmentierung. Es unterstützt bei der Auswahl eines geeigneten Zielpublikums.



Bedarfs- und Absatzprognosen

Data Mining hilft Unternehmen, Verkaufszeiten und Zielgruppen für Produkte besser zu erkennen, indem es Preisoptimierung ermöglicht.

4

Data Mining Methoden

Data Mining

Methoden

Klassifikation auf der Basis von Eigenschaften

- Objekte werden anhand von Ähnlichkeiten oder Mustern in Klassen eingeordnet.
- Anwendung von Techniken wie neuronale Netze, Bayes-Klassifikation, oder Entscheidungsbaum.
- Beispiel: Affinität von Kunden für neue Kampagnen wird als Wahrscheinlichkeit berechnet.
- Wiederholung des Prozesses für mehrere Kampagnen zur optimalen Kundensegmentierung.

Data Mining

Methoden

Gruppierung durch Cluster-Analyse und Segmentierung

- Daten werden in kleinere, homogene Gruppen auf der Basis gemeinsamer Eigenschaften eingeteilt.
- Herausforderung: Algorithmen finden Gruppen ohne Vorwissen, menschliche Analyse zur Erkennung der Ähnlichkeiten erforderlich.
- Ziel ist die Beherrschung großer Datenmengen & Identifizierung verwertbarer Muster.
- Segmentierung hilft bei der zielgerichteten Analyse und der Verbesserung von Geschäftsstrategien.

Data Mining

Methoden

Abhängigkeiten durch Assoziationsregeln und Sequenzen entdecken

- Aufdeckung von Zusammenhängen durch Analyse wiederkehrender Abfolgen und Assoziationsregeln.
- Ziel: Erkennung und Nutzung häufiger Kombinationen oder Muster.
- Anwendung im E-Commerce zur Identifizierung von Regelmäßigkeiten im Kundenverhalten.
- Beispiel: Analyse von Kundenreaktionen auf Website-Ladezeiten zur Optimierung der User Experience.

5

Data Mining

Darauf solltest du achten

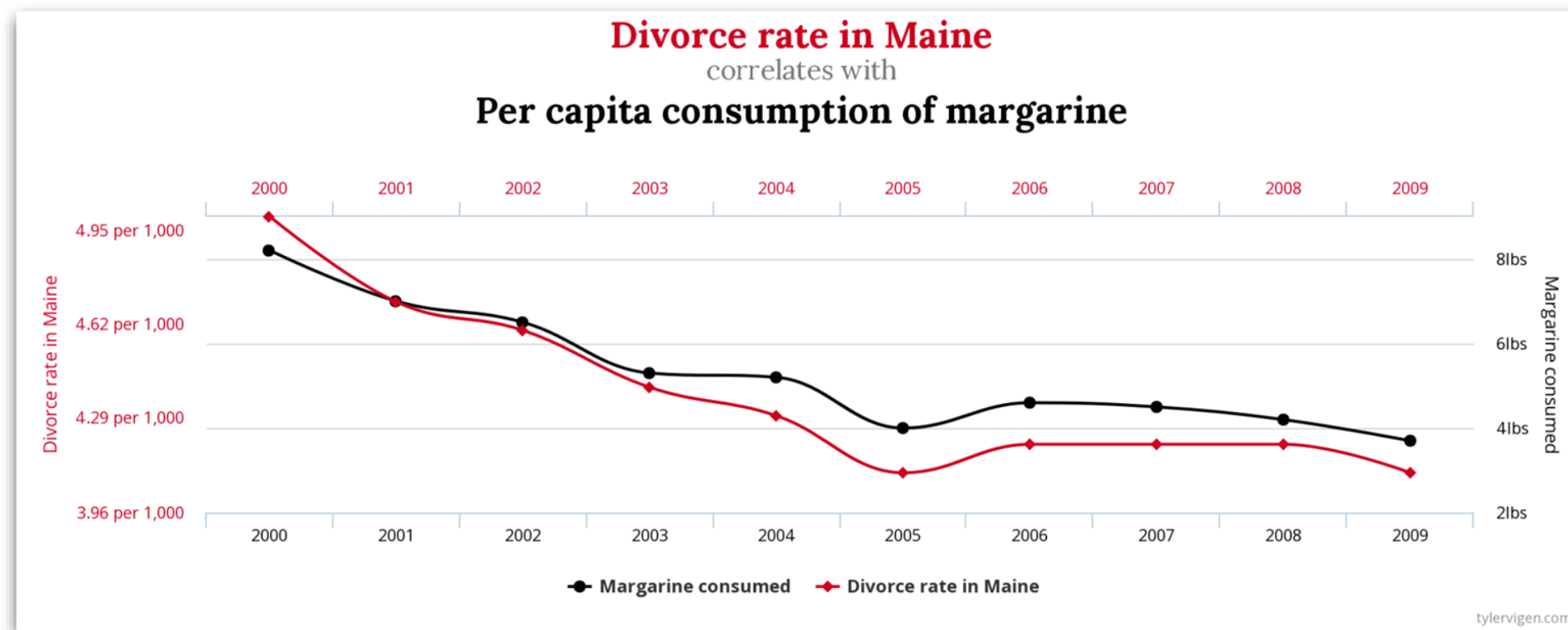
Data Mining

Darauf solltest du achten

- Der Prozess und die Ergebnisse sind unvorhersehbar und lässt sich teilweise auch nicht leicht einordnen
- Selbst wenn es eine Ähnlichkeit in den Daten gibt, heißt das nicht zwingend, dass eine Sache die andere beeinflusst

Data Mining

Darauf solltest du achten



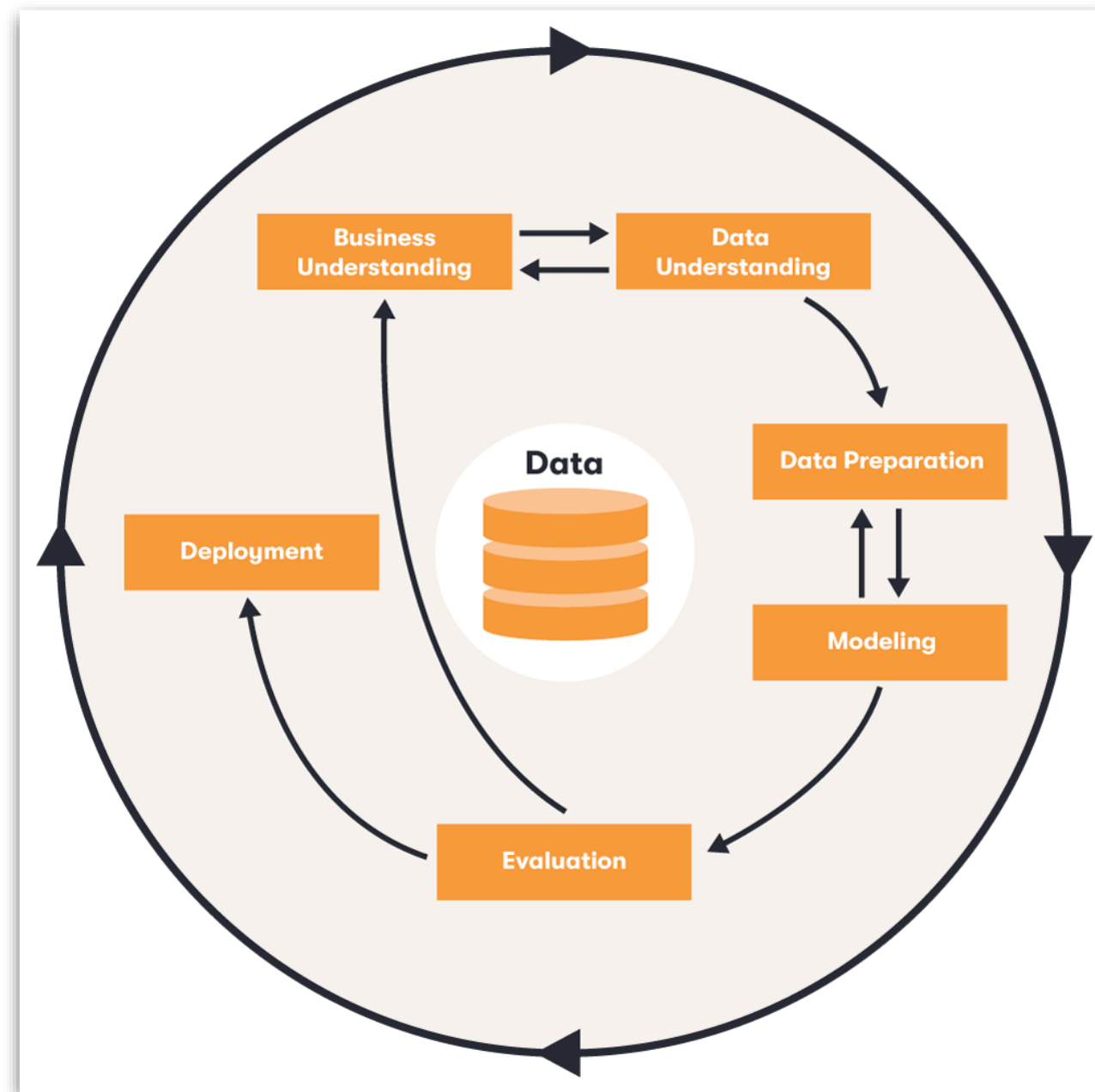
- Das Diagramm zum Verzehr von Margarine ähnelt stark dem Diagramm, das die Scheidungsrate in Maine beschreibt
- Mehr Konsum von Margarine führt zu mehr Scheidungen?
- Vorsicht, wie du die Ergebnisse interpretierst!

5

Data Mining CRISP-DM Workflow

Data Mining

CRISP-DM Workflow



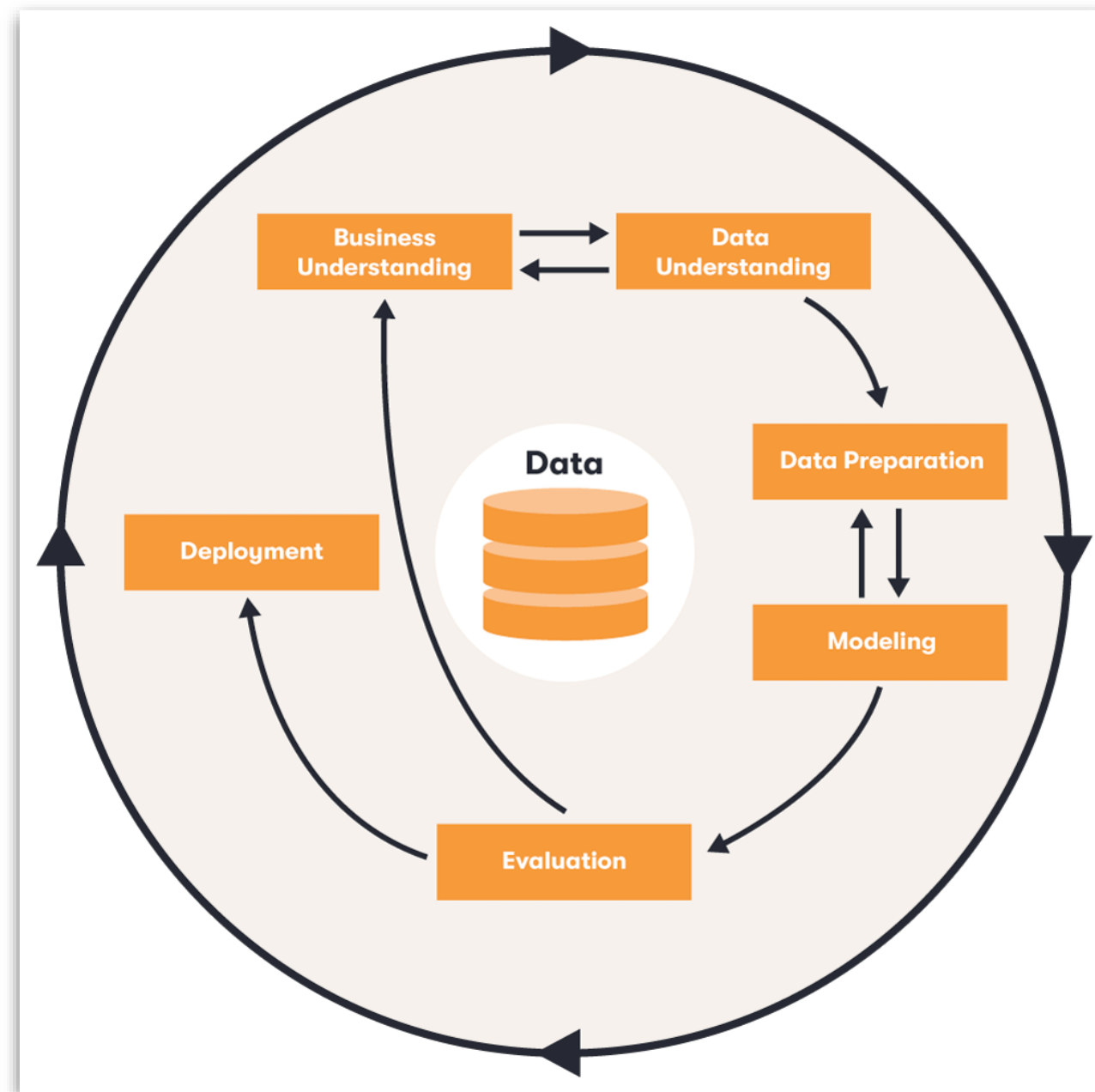
Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Business understanding

- Definition der Geschäftsziele und Einschätzung des aktuellen Zustands.
- Klärung der Anforderungen und Rahmenbedingungen.
- Festlegung der Erfolgskriterien und Durchführung einer Kosten-Nutzen-Bewertung.

Data Mining

CRISP-DM Workflow



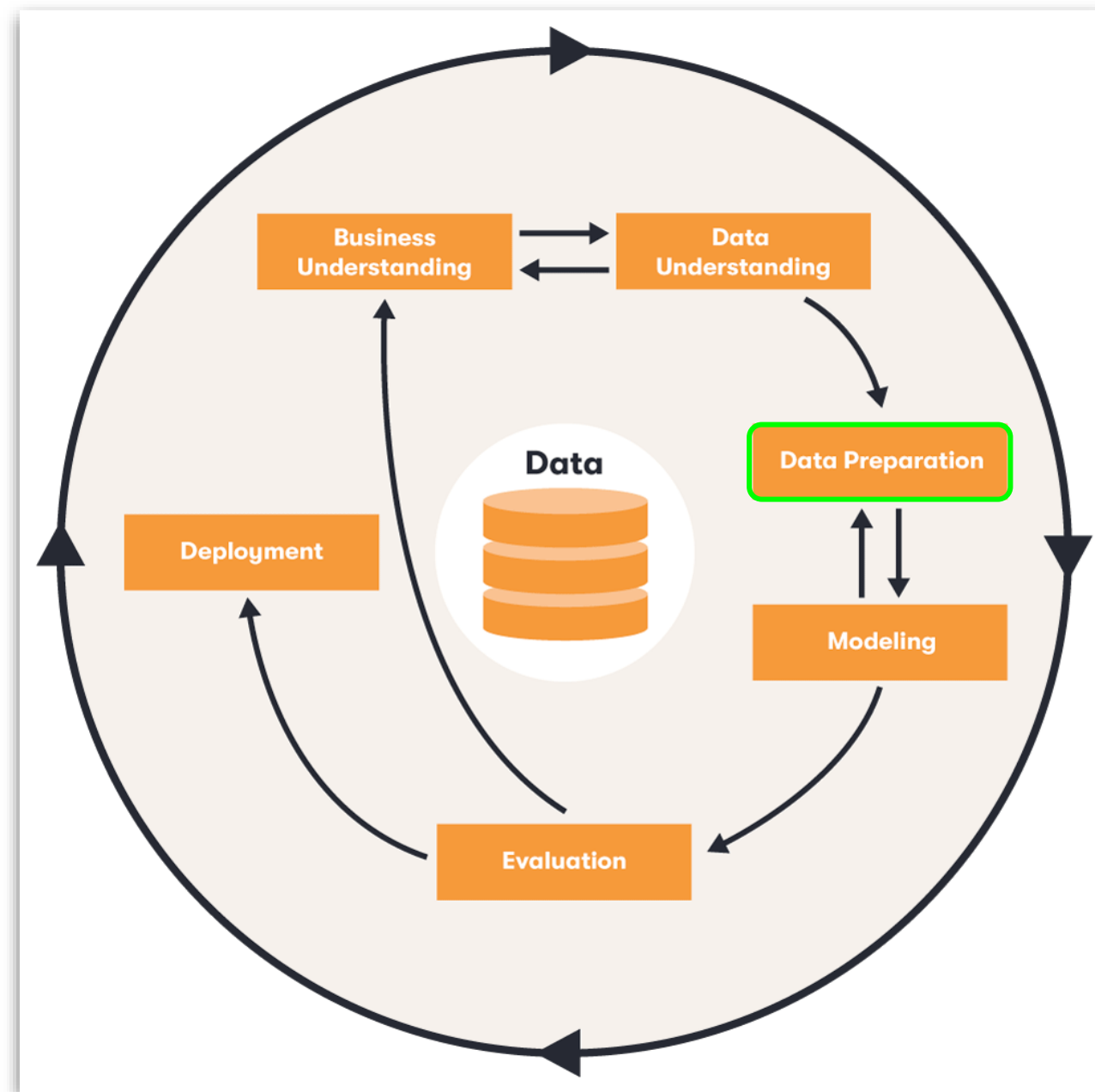
Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Data understanding

- Sammlung der verfügbaren Daten und Identifizierung relevanter Datenquellen.
- Beurteilung der Datenqualität und Identifizierung möglicher Probleme.
- Explorative Datenanalyse, um Einsichten zu gewinnen und Hypothesen zu entwickeln.

Data Mining

CRISP-DM Workflow



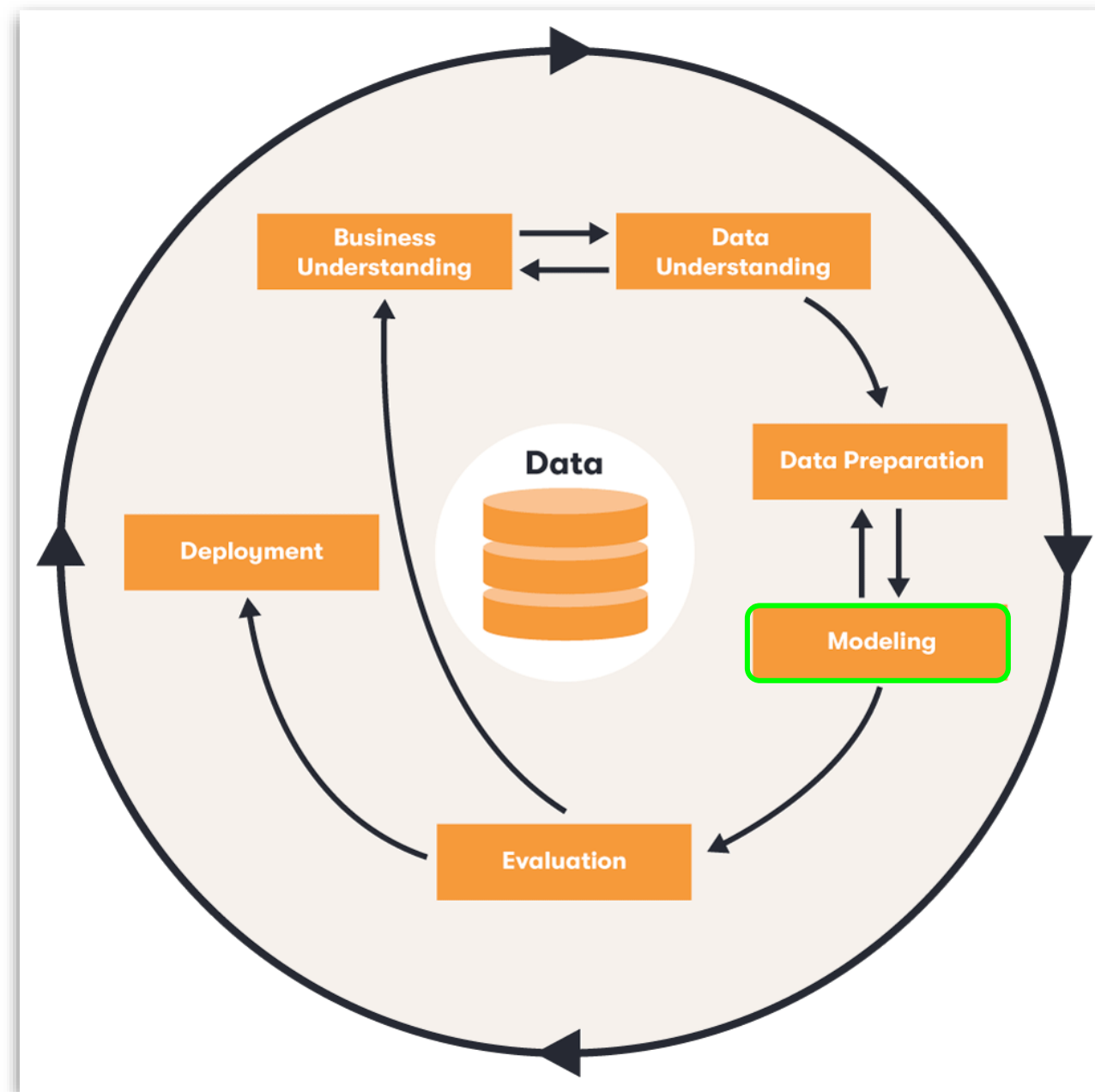
Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Data Preparation

- Auswahl der relevanten Daten und Ausschluss irrelevanter Daten.
- Reinigung der Daten von Fehlern, Ausreißern und Inkonsistenzen.
- Transformation der Daten in geeignete Formate und Strukturen für die Modellierung.

Data Mining

CRISP-DM Workflow



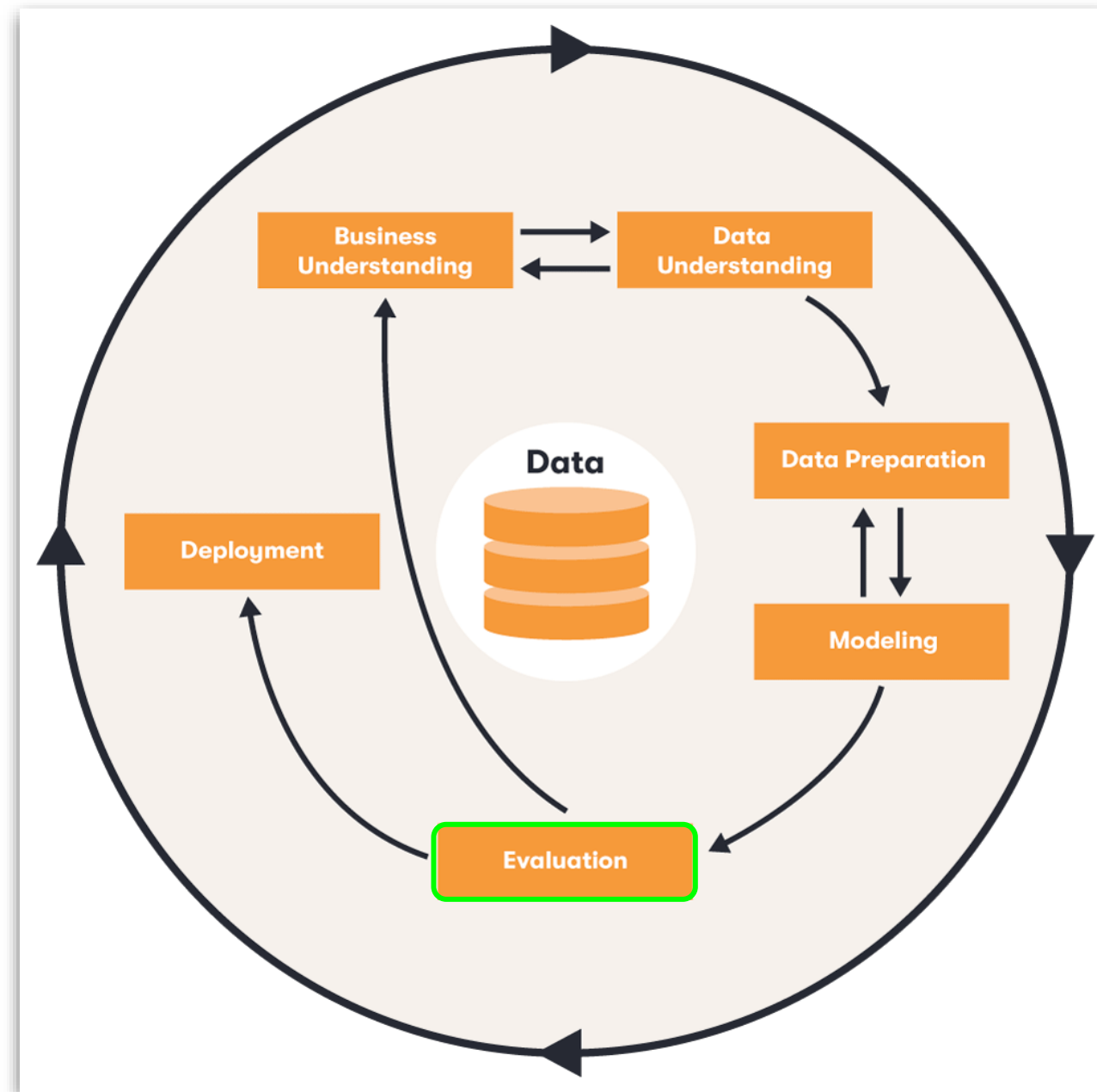
Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Modeling

- Auswahl geeigneter Data-Mining-Methoden und -Algorithmen.
- Planung der Tests und Kreuzvalidierung zur Modellbewertung.
- Entwicklung von mehreren Modellen und Anpassung der Parameter.

Data Mining

CRISP-DM Workflow



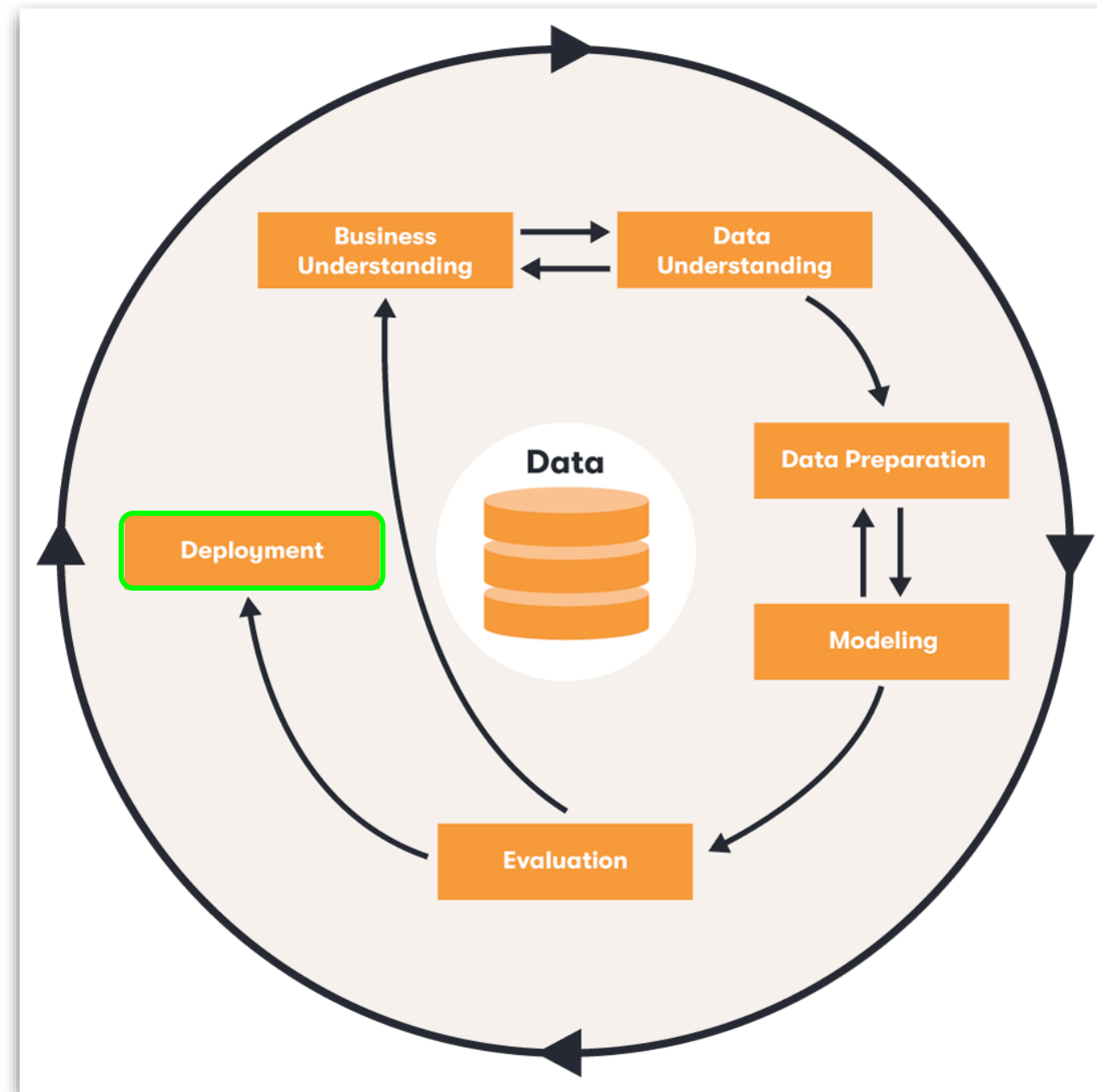
Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Evaluation

- Bewertung der Modelle auf Basis der vorher festgelegten Erfolgskriterien.
- Überprüfung der Tauglichkeit der Modelle für die Geschäftsziele.
- Gegebenenfalls Anpassung und Optimierung der Modelle.

Data Mining

CRISP-DM Workflow



Quelle: <https://www.abtasty.com/wp-content/uploads/data-mining-crisp-dm-modell-ab-tasty.png>

Deployment

- Planung und Vorbereitung der Implementierung der Modelle in die Betriebsumgebung.
- Umsetzung der Modelle und Integration in Geschäftsprozesse.
- Überwachung der Modell-Leistung und Anpassung bei Veränderungen.

6

Data Mining Tools & Kosten

Data Mining

Tools & Kosten

- Es gibt ein paar nützliche Tools, die beim Data Mining unterstützen und die manuelle Arbeit übernehmen können.
- Viele Tools sind jedoch ziemlich teuer. Du kannst die Daten auch in eine Excel übertragen und selbst auswerten, Tools sind aber einfacher und effizienter
- Wäge ab, ob du ein Budget für zeitsparende Data Mining Tools investieren willst.
- Viele Tools bieten eine kostenlose Testphase an, andere sind Open-Source.

Data Mining

Tools & Kosten

The screenshot shows the Oracle Cloud Infrastructure (OCI) website. The top navigation bar includes the OCI logo, links for About, Services, Solutions, Pricing, Partners, and Resources, a search icon, a share icon, and a 'Sign in to Oracle Cloud' button. Below the navigation bar, there's a 'Database >' breadcrumb and a 'Try Oracle Cloud Free Tier' button. The main heading is 'Oracle Database Technologies'. The text describes Oracle Database's performance, scalability, reliability, and security, mentioning Oracle Database 19c as the current 'long term release' and Oracle Database 21c as an 'innovation release'. A video player shows a man speaking, with the caption 'Oracle Database Free for developers (3:40)'. At the bottom, there's a section titled 'Announcing Oracle Database@Azure' with a description of the partnership and two buttons: 'Learn more' and 'Read the press release'.

OCI About Services Solutions Pricing Partners Resources

Database > Try Oracle Cloud Free Tier

Oracle Database Technologies

Oracle Database offers market-leading performance, scalability, reliability, and security, both on-premises and in the cloud. Oracle Database 19c is the current [long term release](#), and it provides the highest level of release stability and longest time-frame for support and bug fixes.

Oracle Database 21c, also available for production use today as an [innovation release](#), provides an early insight into the many enhancements and new capabilities. These include autonomous administration, improved multimodel support through in-database Javascript and native blockchain tables, and multiworkload improvements such as AutoML and sharding enhancements that will be incorporated into future long term releases.

[Read and sign up for Oracle Database newsletters](#)

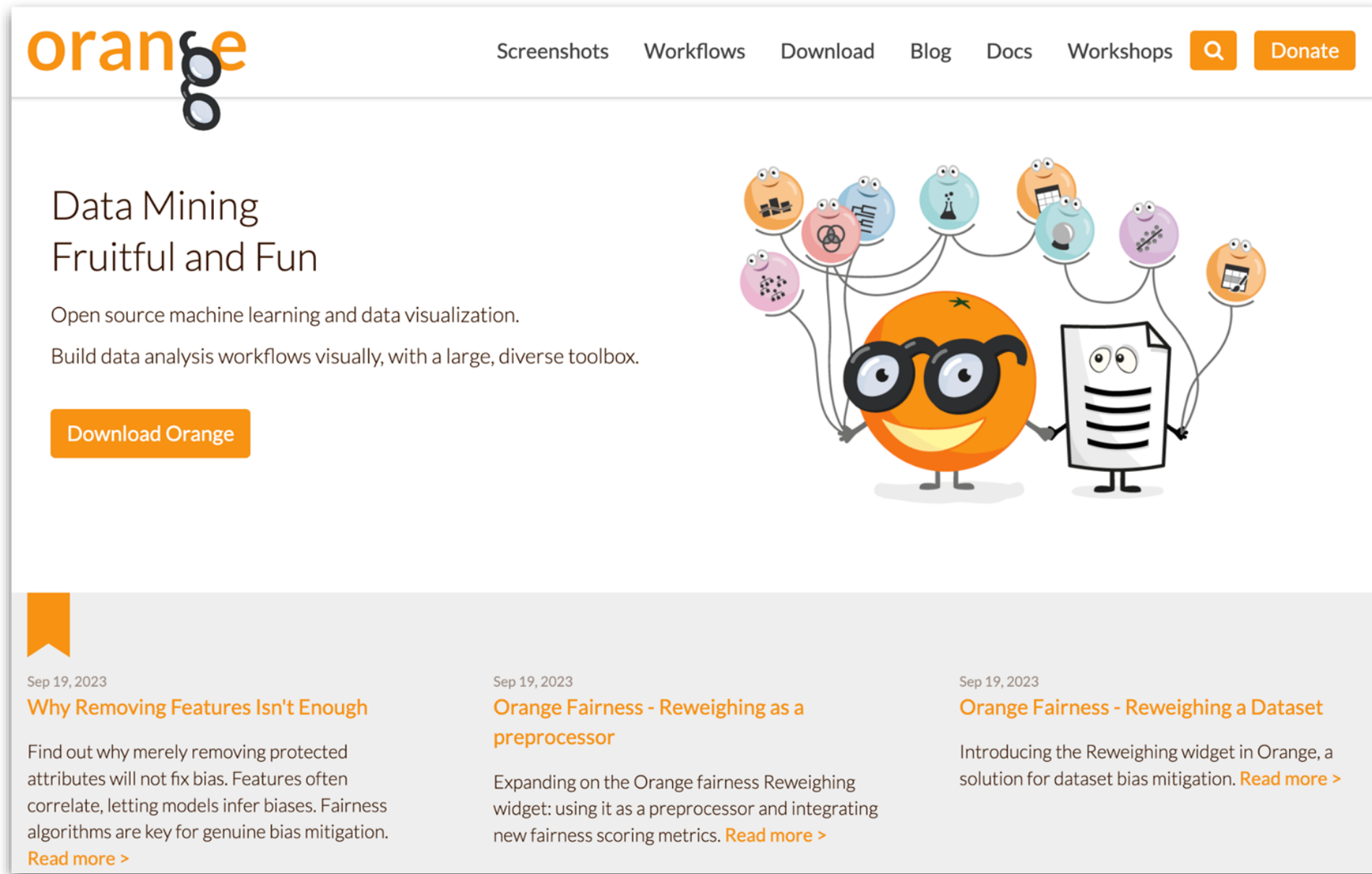
Announcing Oracle Database@Azure

The expanded Oracle-Microsoft partnership, announced by Larry Ellison and Satya Nadella, delivers Oracle database services on Oracle Cloud Infrastructure in Microsoft Azure.

[Learn more](#) [Read the press release](#)

- Gehört zu Oracle Advanced Analytics
- Bietet einige Algorithmen für Klassifikation, Regression, Entdeckung von Anomalien, Vorhersagen, etc.
- 30 Tage kostenlos, dann ab 50€ - 1800€ Monatlich

Data Mining Tools & Kosten



The screenshot shows the Orange Data Mining website. At the top is the 'orange' logo in orange and grey. To its right is a navigation bar with links: 'Screenshots', 'Workflows', 'Download', 'Blog', 'Docs', 'Workshops', a search icon, and a 'Donate' button. Below the navigation bar is a large illustration of a smiling orange character with glasses, holding hands with several smaller, colorful characters (pink, blue, green, orange, purple) that represent different data mining concepts like classification, regression, clustering, etc. To the left of this illustration, the text reads 'Data Mining Fruitful and Fun', followed by 'Open source machine learning and data visualization.' and 'Build data analysis workflows visually, with a large, diverse toolbox.' Below this text is an orange 'Download Orange' button. At the bottom of the page, there is a grey section with three blog posts, each dated 'Sep 19, 2023'. The first post is 'Why Removing Features Isn't Enough' with a brief description and a 'Read more >' link. The second post is 'Orange Fairness - Reweighing as a preprocessor' with a brief description and a 'Read more >' link. The third post is 'Orange Fairness - Reweighing a Dataset' with a brief description and a 'Read more >' link.

orange

Screenshots Workflows Download Blog Docs Workshops [Donate](#)

Data Mining Fruitful and Fun

Open source machine learning and data visualization.
Build data analysis workflows visually, with a large, diverse toolbox.

[Download Orange](#)

Sep 19, 2023
Why Removing Features Isn't Enough
Find out why merely removing protected attributes will not fix bias. Features often correlate, letting models infer biases. Fairness algorithms are key for genuine bias mitigation.
[Read more >](#)

Sep 19, 2023
Orange Fairness - Reweighing as a preprocessor
Expanding on the Orange fairness Reweighing widget: using it as a preprocessor and integrating new fairness scoring metrics. [Read more >](#)

Sep 19, 2023
Orange Fairness - Reweighing a Dataset
Introducing the Reweighing widget in Orange, a solution for dataset bias mitigation. [Read more >](#)

- Orange Data Mining
- Open Source - kostenlos
- Bietet umfangreiche Data Mining Funktionalitäten
- Nutzung kann evtl. komplexer sein