

Umsatz- und Länderanalyse eines internationalen Online-Shops – Datenintegration & Reporting mit SAP Datasphere und SAP Analytics Cloud

Hausarbeit

im Rahmen des Seminars „Business Intelligence 2“

Sommersemester 2025

Daten ausgeblendet für Github
Abgabetermin: 2025-07-21

Inhaltsverzeichnis

1	Einleitung	1
2	Zielsetzung & Leitfragen	2
2.1	Zielsetzung.....	2
2.2	Leitfragen	2
3	Datenbasis & Angedachtes dimensionales Modell.....	3
4	Methodisches Vorgehen.....	5
4.1	Datenimport & Staging	5
4.2	Staging-Area	8
4.3	Modellierung der Dimensionen	9
4.4	Fakten-View	11
4.5	Kennzahlen & Filter-Variablen	12
4.6	Story-Erstellung in SAP Analytics Cloud.....	14
5	Fazit.....	17

1 Einleitung

Der Online-Handel wächst kontinuierlich und erzeugt täglich riesige Mengen an Transaktionsdaten. Damit Unternehmen aus diesen Rohdaten zeitnah belastbare Entscheidungen ableiten können, benötigen sie eine skalierbare, cloudbasierte Analyseumgebung. Herkömmliche On-Premise-Data-Warehouse-Lösungen stoßen dabei häufig an Grenzen: Sie sind schwerfällig zu erweitern, benötigen hohe Administrationsaufwände und binden Datenquellen oft nur mit erheblichem Verzögerungsaufwand ein.¹

In dieser Arbeit wird daher exemplarisch gezeigt, wie ein frei verfügbarer E-Commerce-Datensatz aus Kaggle („E-commerce Data“) mithilfe von **SAP Datasphere** in ein konsistentes Cloud-Data-Warehouse überführt und anschließend in **SAP Analytics Cloud (SAC)** interaktiv ausgewertet werden kann. Nach der Datenbereinigung und dem Upload werden dafür zunächst zentrale Dimensionen wie Produkt, Kunde und Zeit modelliert und mit den Verkaufsfakten verknüpft.² Auf Basis dieses Analytic Models werden anschließend Kennzahlen wie Gesamtumsatz, durchschnittlicher Warenkorbwert sowie ein Länder- und Zeitvergleich abgeleitet und in einem SAC-Dashboard visualisiert.

Ziel der Dokumentation ist es, den gesamten End-to-End-Prozess – von der Rohdatei bis zum fertig nutzbaren Management-Dashboard – transparent nachzuvollziehen und gleichzeitig den Mehrwert eines modernen Cloud-BI-Stacks für mittelständische E-Commerce-Anbieter herauszuarbeiten. Die Arbeit demonstriert damit praxisnah, wie auch nicht-SAP-Quellen nahtlos in die SAP-Analyselandschaft integriert werden können und welche Erkenntnisse sich aus einer konsistenten Datenbasis gewinnen lassen.

¹ Corporation (2024).

² Vgl. Kimball/Ross (2002), S. 21.

2 Zielsetzung & Leitfragen

2.1 Zielsetzung

Diese Hausarbeit verfolgt das Ziel, einen durchgängigen End-to-End-Prozess zur Integration und Analyse von Online-Shop-Verkaufsdaten mit SAP Datasphere und SAP Analytics Cloud zu dokumentieren. Ausgangspunkt ist der frei verfügbare Kaggle-Datensatz „E-commerce Data“, der zunächst bereinigt und in SAP Datasphere geladen wird.³ Anschließend werden auf Basis eines dimensional Modells zentrale Kennzahlen (u. a. Gesamtumsatz, verkaufte Stückzahlen, durchschnittlicher Warenkorbwert) definiert, bevor sie in einem interaktiven SAC-Dashboard visualisiert und interpretiert werden. Die Arbeit soll zeigen, wie sich nicht-SAP-Quellen mit den Cloud-BI-Werkzeugen von SAP konsistent integrieren lassen und welchen geschäftlichen Mehrwert eine derart schnelle Datenaufbereitung für mittelständische E-Commerce-Anbieter bietet.

2.2 Leitfragen

Datenintegration

Wie lassen sich CSV-Verkaufsdaten in SAP Datasphere so aufbereiten, dass ein harmonisiertes, dimensionsorientiertes Datenmodell entsteht?

Kennzahlendefinition

Welche Kennzahlen eignen sich, um Umsatz, Bestellverhalten und Länderunterschiede eines Online-Shops aussagekräftig abzubilden?

Geschäftlicher Mehrwert

Welche konkreten Erkenntnisse lassen sich aus dem integrierten Dashboard ableiten, und wie unterstützen sie strategische Entscheidungen?

³ Carrie (2017).

3 Datenbasis & Angedachtes dimensionales Modell

Als empirische Grundlage dient der öffentlich verfügbare Kaggle-Datensatz „**E-commerce Data**“.⁴ Er enthält rund 540 000 Transaktionszeilen eines britischen Online-Händlers aus den Jahren 2010–2011 und weist acht Kernspalten auf: *InvoiceNo* (Bestellnummer), *StockCode* (Artikelcode), *Description* (Artikeltext), *Quantity*, *InvoiceDate*, *UnitPrice*, *CustomerID* und *Country*. Die Daten decken damit den gesamten Bestellprozess ab – vom einzelnen Artikel bis zur zugehörigen Rechnungs- und Kundeninformation – und sind länderübergreifend (UK sowie über 30 Exportländer) erhoben.

Für die Modellierung in SAP Datasphere wird der Roh-CSV-Export zunächst lokal geprüft und bereinigt:

- **Stornobelege** (*InvoiceNo* beginnt mit „C“) werden entfernt, ebenso Datensätze mit $Quantity \leq 0$.
- **Fehlende Kundennummern** erhalten den Platzhalterwert „Unknown“, um die Integrität der Customer-Dimension zu wahren.
- Das Zeitfeld *InvoiceDate* wird aus dem Textformat in den SQL-Date-Typ überführt.
- Per Berechnungsspalte $TotalPrice = Quantity \times UnitPrice$ wird der Zeilenumsatz ergänzt.

Das Datenmodell besteht aus fünf Dimensionstabellen – Product, Customer, Country, Invoice und Date – sowie dem zentralen Fact-View 267_FACT_Sales.

Die eigenständige Country-Dimension erlaubt präzise Länderanalysen (etwa UK vs. Non-UK), während die Invoice-Dimension Distinct-Kennzahlen wie Bestellfrequenz und durchschnittlichen Warenkorb performant unterstützt. Dank der klaren Sternstruktur, des hohen Transaktionsvolumens und der freien Lizenz eignet sich der Datensatz optimal, um eine End-to-End-Integration in der SAP-Cloud-BI-Umgebung zu demonstrieren.

⁴ Carrie (2017).

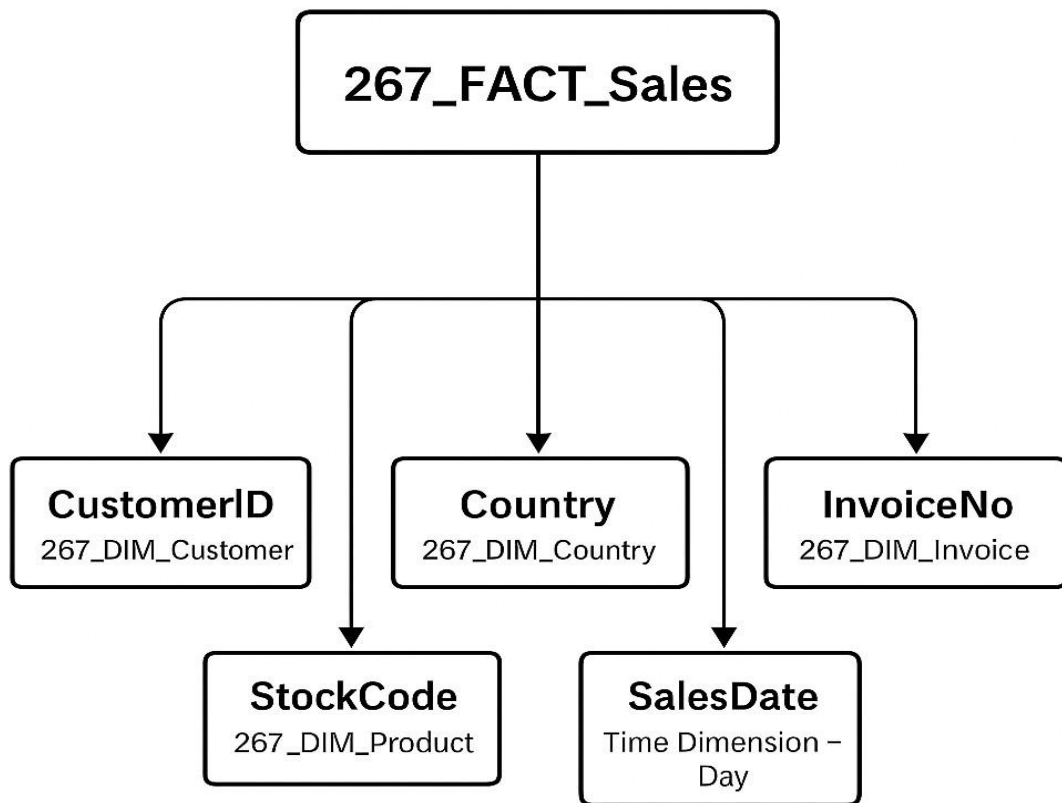


ABBILDUNG 1: ANGEDACHTES DATENMODELL⁵

⁵ Abdulhadi/Github (2025).

4 Methodisches Vorgehen

4.1 Datenimport & Staging

Nach einer ersten Sichtung des Kaggle-CSV-Exports werden offensichtliche Datenqualitätsprobleme (Stornobelege, negative Mengen, fehlende Kundennummern) in Excel bzw. Python bereinigt.⁶

Folgende Abbildungen werden in Excel – PowerQuery abgebildet, um dem Leser eine bessere Übersicht zu bieten.

StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
85123A	WHITE HANGING HEART T-LIGHT HOLD...	6	12.01.2010 08:24			United Kingdom
71053	WHITE METAL LANTERN	6	12.01.2010 08:26			United Kingdom
84406B	CREAM CUPID HEARTS COAT HANGER	8	12.01.2010 08:26			United Kingdom
84029G	KNITTED UNION FLAG HOT WATER BOT...	6	12.01.2010 08:26			United Kingdom
84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12.01.2010 08:26			United Kingdom
22752	SET 7 BABUSHKA NESTING BOXES	2	12.01.2010 08:26			United Kingdom
21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	12.01.2010 08:26			United Kingdom
22633	HAND WARMER UNION JACK	6	12.01.2010 08:26			United Kingdom
22632	HAND WARMER RED POLKA DOT	6	12.01.2010 08:26			United Kingdom
84879	ASSORTED COLOUR BIRD ORNAMENT	32	12.01.2010 08:34			United Kingdom
22745	POPPY'S PLAYHOUSE BEDROOM	6	12.01.2010 08:34			United Kingdom
22748	POPPY'S PLAYHOUSE KITCHEN	6	12.01.2010 08:34			United Kingdom
22749	FELTCRAFT PRINCESS CHARLOTTE DOLL	8	12.01.2010 08:34			United Kingdom
22310	IVORY KNITTED MUG COSY	6	12.01.2010 08:34			United Kingdom
84969	BOX OF 6 ASSORTED COLOUR TEASPOO...	6	12.01.2010 08:34			United Kingdom
22623	BOX OF VINTAGE JIGSAW BLOCKS	3	12.01.2010 08:34			United Kingdom
22622	BOX OF VINTAGE ALPHABET BLOCKS	2	12.01.2010 08:34			United Kingdom
21754	HOME BUILDING BLOCK WORD	3	12.01.2010 08:34			United Kingdom
21755	LOVE BUILDING BLOCK WORD	3	12.01.2010 08:34			United Kingdom
21777	RECIPE BOX WITH METAL HEART	4	12.01.2010 08:34			United Kingdom
48187	DOORMAT NEW ENGLAND	4	12.01.2010 08:34			United Kingdom
22960	JAM MAKING SET WITH JARS	6	12.01.2010 08:34			United Kingdom
22913	RED COAT RACK PARIS FASHION	3	12.01.2010 08:34			United Kingdom
22912	YELLOW COAT RACK PARIS FASHION	3	12.01.2010 08:34			United Kingdom
22914	BLUE COAT RACK PARIS FASHION	3	12.01.2010 08:34			United Kingdom
21756	BATH BUILDING BLOCK WORD	3	12.01.2010 08:35			United Kingdom
22728	ALARM CLOCK BAKELIKE PINK	24	12.01.2010 08:43			France
22727	ALARM CLOCK BAKELIKE RED	24	12.01.2010 08:43			France
22726	ALARM CLOCK BAKELIKE GREEN	12	12.01.2010 08:43			France

ABBILDUNG 2: DATEN UNBEARBEITET – POWERQUERY

Zur besseren Übersicht wurden aus den Datensätzen folgende Bestellnummern ausgewählt, um die Datenbereinigung nachvollziehbar darzustellen.

- 536589
- 536402
- C571196

⁶ Abdulhadi/Github (2025).

Zeilen filtern

Hiermit wenden Sie mindestens eine Filterbedingung auf die Zeilen in dieser Tabelle an.

☐ Standard ☒ Weitere

Zeilen beibehalten, für die Folgendes gilt:

Und/oder	Spalte	Operator	Wert
	InvoiceNo	ist gleich	536589
Od...	InvoiceNo	ist gleich	536402
Od...	InvoiceNo	ist gleich	C571196

Klausel hinzufügen

OK Abbrechen

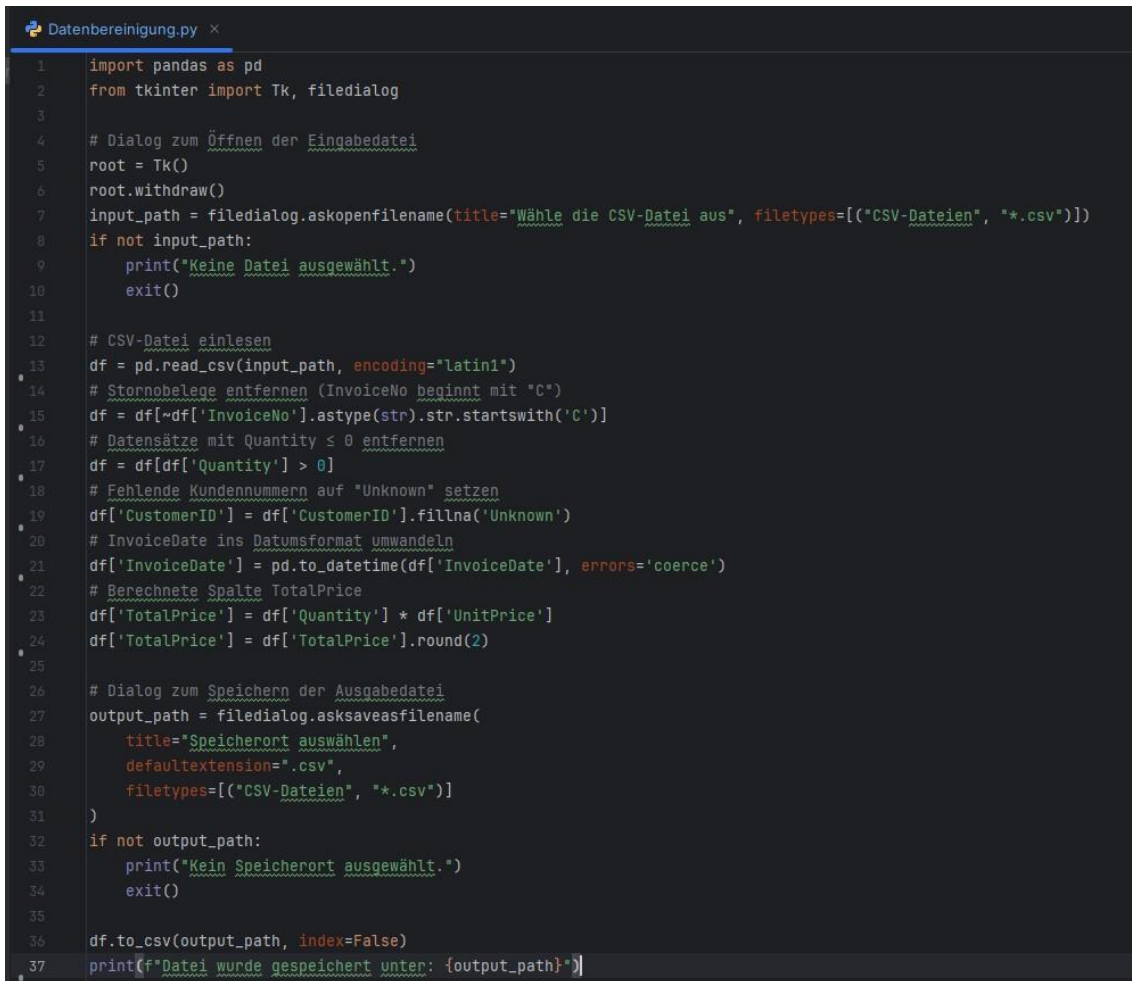
ABBILDUNG 3: FILTER - BEISPIELE

Abbildung 3 zeigt den entsprechenden Filter, der gesetzt wurde, um die nötigen Bestellungen anzuzeigen.

	A	B	C	D	E	F	G	H
1	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
2	536402	22086	PAPER CHAIN KIT 50'S CHRISTMAS	40	12.01.2010 11:22	255	15513	United Kingdom
3	536402	22910	PAPER CHAIN KIT VINTAGE CHRISTMAS	40	12.01.2010 11:22	255	15513	United Kingdom
4	536402	22837	HOT WATER BOTTLE BABUSHKA	36	12.01.2010 11:22	425	15513	United Kingdom
5	536589	21777		-10	12.01.2010 16:50	0		United Kingdom
6	C571196	84991	60 TEATIME FAIRY CAKE CASES	-24		55		United Kingdom
7	C571196	23483	HANGING BUTTERFLY T-LIGHT HOLDER	-12		125		United Kingdom
8	C571196	23354	6 GIFT TAGS 50'S CHRISTMAS	-12		83		United Kingdom
9	C571196	23353	6 GIFT TAGS VINTAGE CHRISTMAS	-12		83		United Kingdom
10	C571196	23352	ROLL WRAP 50'S RED CHRISTMAS	-12		125		United Kingdom
11	C571196	23351	ROLL WRAP 50'S CHRISTMAS	-12		125		United Kingdom
12	C571196	23350	ROLL WRAP VINTAGE SPOT	-12		125		United Kingdom
13	C571196	23349	ROLL WRAP VINTAGE CHRISTMAS	-12		125		United Kingdom
14	C571196	23311	VINTAGE CHRISTMAS STOCKING	-6		255		United Kingdom
15	C571196	23225	CHERUB HEART DECORATION SILVER	-12		83		United Kingdom
16	C571196	23223	CHRISTMAS TREE HANGING SILVER	-12		83		United Kingdom
17	C571196	23222	CHRISTMAS TREE HANGING GOLD	-12		83		United Kingdom
18	C571196	23220	REINDEER HEART DECORATION GOLD	-12		83		United Kingdom
19	C571196	22866	HAND WARMER SCOTTY DOG DESIGN	-12		21		United Kingdom
20	C571196	22809	SET OF 6 T-LIGHTS SANTA	-6		295		United Kingdom
21	C571196	22632	HAND WARMER RED RETROSPOT	-12		21		United Kingdom
22	C571196	22383	LUNCH BAG SUKI DESIGN	-10		165		United Kingdom
23	C571196	21731	RED TOADSTOOL LED NIGHT LIGHT	-12		165		United Kingdom
24	C571196	21212	PACK OF 72 RETROSPOT CAKE CASES	-24		55		United Kingdom
25	C571196	20727	LUNCH BAG BLACK SKULL	-10		165		United Kingdom
26	C571196	20725	LUNCH BAG RED RETROSPOT	-10		165		United Kingdom
27								
28								

ABBILDUNG 4: UNBEREINIGTE DATEN MIT FILTER – EXCEL

Abbildung 4 zeigt stornierte Belege, fehlende CustomerIDs und stornierte Bestellungen. Diese müssen bereinigt werden, da stornierte Bestellungen für die benötigten Kennzahlen nicht relevant sind.



```

1 import pandas as pd
2 from tkinter import Tk, filedialog
3
4 # Dialog zum Öffnen der Eingabedatei
5 root = Tk()
6 root.withdraw()
7 input_path = filedialog.askopenfilename(title="Wähle die CSV-Datei aus", filetypes=[("CSV-Dateien", "*.csv")])
8 if not input_path:
9     print("Keine Datei ausgewählt.")
10    exit()
11
12 # CSV-Datei einlesen
13 df = pd.read_csv(input_path, encoding="latin1")
14 # Stornobelege entfernen (InvoiceNo beginnt mit "C")
15 df = df[~df['InvoiceNo'].astype(str).str.startswith('C')]
16 # Datensätze mit Quantity ≤ 0 entfernen
17 df = df[df['Quantity'] > 0]
18 # Fehlende Kundennummern auf "Unknown" setzen
19 df['CustomerID'] = df['CustomerID'].fillna('Unknown')
20 # InvoiceDate ins Datumsformat umwandeln
21 df['InvoiceDate'] = pd.to_datetime(df['InvoiceDate'], errors='coerce')
22 # Berechnete Spalte TotalPrice
23 df['TotalPrice'] = df['Quantity'] * df['UnitPrice']
24 df['TotalPrice'] = df['TotalPrice'].round(2)
25
26 # Dialog zum Speichern der Ausgabedatei
27 output_path = filedialog.asksaveasfilename(
28     title="Speicherort auswählen",
29     defaultextension=".csv",
30     filetypes=[("CSV-Dateien", "*.csv")]
31 )
32 if not output_path:
33     print("Kein Speicherort ausgewählt.")
34     exit()
35
36 df.to_csv(output_path, index=False)
37 print(f"Datei wurde gespeichert unter: {output_path}")

```

ABBILDUNG 5: DATENBEREINIGUNGSSKRIPT - PYTHON⁷

Abbildung 5 zeigt ein Python-Skript, das zur Bereinigung von Verkaufsdaten im CSV-Format dient. Es ermöglicht die interaktive Auswahl einer Eingabedatei und eines Speicherorts für die bereinigte Ausgabedatei. Im Skript werden unter anderem folgende Schritte durchgeführt: Entfernung stornierter Belege (beginnend mit „C“), Ausschluss von Datensätzen mit nicht-positiver Menge, Ersetzung fehlender Kundennummern durch „Unknown“, Umwandlung des Datumsformats sowie Berechnung eines neuen Attributs „TotalPrice“. Das Skript automatisiert so zentrale Bereinigungs-schritte für eine saubere Datenbasis.

⁷ Abdulhadi/Github (2025).

	A	B	C	D	E	F	G	H	I	J
1	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country	TotalPrice	
375	536402	22086	PAPER CHAIN KIT 50'S CHRISTMAS	40	01.12.2010 11:22	255	155130	United Kingdom	1020	
376	536402	22910	PAPER CHAIN KIT VINTAGE CHRISTMAS	40	01.12.2010 11:22	255	155130	United Kingdom	1020	
377	536402	22837	HOT WATER BOTTLE BABUSHKA	36	01.12.2010 11:22	425	155130	United Kingdom	1530	
331287										
331288										
331289										
331290										
331291										
331292										
331293										
331294										

ABBILDUNG 6: BEREINIGTE DATEN MIT FILTER - EXCEL

- Die bereinigte Datei **daten_bereinigt.csv** wird anschließend in den zugeordneten Datasphere-Space hochgeladen.
- Während des Imports werden Datentypen vorerst als „String“ belassen, um spätere Typ-Konvertierungen zentral im Data Builder vorzunehmen.

4.2 Staging-Area

Um die Tabelle in das System von Datasphere einbinden zu können und die Daten weiterverarbeiten zu können, musste die Tabelle zunächst importiert werden.

	22 InvoiceNo	AA StockCode	AA Description	22 Quantity	AA InvoiceDate	123 UnitPrice	AA CustomerID	AA Country	123 TotalPrice
551	541221	22631	CIRCUS PARADE LU	2	2011-01-14 14:28:00	3.29	Unknown	United Kingdom	6.58
552	541221	22637	PIGGY BANK RETRC	4	2011-01-14 14:28:00	5.79	Unknown	United Kingdom	23.16
553	541221	22644	CERAMIC CHERRY C	1	2011-01-14 14:28:00	4.13	Unknown	United Kingdom	4.13
554	541221	22646	CERAMIC STRAWBE	4	2011-01-14 14:28:00	4.13	Unknown	United Kingdom	16.52
555	541221	22652	TRAVEL SEWING KIT	5	2011-01-14 14:28:00	4.13	Unknown	United Kingdom	20.65
556	541221	22654	DELUXE SEWING KIT	4	2011-01-14 14:28:00	12.46	Unknown	United Kingdom	49.84
557	541221	22694	WICKER STAR	2	2011-01-14 14:28:00	3.29	Unknown	United Kingdom	6.58
558	541221	22728	ALARM CLOCK BAK	3	2011-01-14 14:28:00	5.79	Unknown	United Kingdom	17.37
559	541221	22738	RIBBON REEL SNOV	2	2011-01-14 14:28:00	2.46	Unknown	United Kingdom	4.92
560	541221	22739	RIBBON REEL CHR!!	1	2011-01-14 14:28:00	2.46	Unknown	United Kingdom	2.46
561	541221	22835	HOT WATER BOTTLE	2	2011-01-14 14:28:00	7.46	Unknown	United Kingdom	14.92
562	541221	35970	ZINC FOLKART SLEI	3	2011-01-14 14:28:00	4.13	Unknown	United Kingdom	12.39
563	541221	48189	DOORMAT FRIENDS	1	2011-01-14 14:28:00	16.63	Unknown	United Kingdom	16.63
564	541221	82494I	WOODEN FRAME AP	1	2011-01-14 14:28:00	8.29	Unknown	United Kingdom	8.29
565	541221	84029E	RED WOOLLY HOTTI	1	2011-01-14 14:28:00	8.29	Unknown	United Kingdom	8.29
566	541221	84029G	KNITTED UNION FLJ	4	2011-01-14 14:28:00	8.29	Unknown	United Kingdom	33.16
567	541221	84380	SET OF 3 BUTTERFL	1	2011-01-14 14:28:00	3.29	Unknown	United Kingdom	3.29
568	541221	84692	BOX OF 24 COCKTA	9	2011-01-14 14:28:00	1.63	Unknown	United Kingdom	14.67
569	541221	21866	UNION JACK FLAG I	1	2011-01-14 14:28:00	3.29	Unknown	United Kingdom	3.29
570	541221	84991	60 TEATIME FAIRY C	2	2011-01-14 14:28:00	1.25	Unknown	United Kingdom	2.5
571	541221	21929	JUMBO BAG PINK V	18	2011-01-14 14:28:00	4.96	Unknown	United Kingdom	89.28
572	541221	21930	JUMBO STORAGE B	6	2011-01-14 14:28:00	4.13	Unknown	United Kingdom	24.78

ABBILDUNG 7: IMPORTIERTE TABELLE - STAGING E-COMMERCE RAW⁸

Folgende Spalten wurden umbenannt:

- StockCode -> ProductID
- Description -> ProductDesc

⁸ Abdulhadi/Github (2025).

List Objects In: All > 267_ECom_Project > D1_Staging (1)							
<input type="checkbox"/> Business Name	Technical Name	Type (Semantic Usage)	Space	Folder	Status	Changed On	Actions
<input checked="" type="checkbox"/> Staging E-Commerce Raw	Staging_ECommerce_Raw	Local Table (Relational Dataset)	AW/004	D1_Staging	Deployed	Jun 4, 2025, 23:32:45	☆

ABBILDUNG 8: STAGING E-COMMERCE RAW - VERZEICHNIS⁹

4.3 Modellierung der Dimensionen

Product-Dimension (DIM_PRODUCT)

1. Graphical View anlegen, Key = *ProductID*.
2. ProductDesc -> Semantic Type: Text & mit ProductID associated
3. Business & Technical Name: 267_DIM_Product



ABBILDUNG 9: PRODUCT - DIMENSION¹⁰

Customer-Dimension (DIM_CUSTOMER)\

1. *CustomerID* als Key, *Country* als Attribut.
2. Hierarchie „Country > CustomerID“ definieren.

Hierarchy Dialog

CustHier

CustHier

Level-Based Hierarchy

+

▼

🗑️

Business Name:

CustHier

Technical Name:

CustHier

Levels

Country

▼

✕

CustomerID

▼

✕

+

Close

ABBILDUNG 10: HIERARCHIE-DIALOG COUNTRY-DIMENSION¹¹

⁹ Abdulhadi/*Github* (2025).

¹⁰ Abdulhadi/*Github* (2025).

¹¹ Abdulhadi/*Github* (2025).

Country-Dimension (DIM_COUNTRY)

1. *SQL View* → Projection nur auf Country.
2. SQL Code eingeben:

```
SELECT DISTINCT "Country"
FROM "Staging_ECommerce_Raw"
WHERE "Country" IS NOT NULL
```

3. Key = Country.
4. Semantic Type = Text

Durch das SELECT DISTINCT erhält man jede Länder-ausprägung genau einmal. Country bleibt als Schlüssel-attribut der Dimension und kann später in Zeit-, Produkt- oder Kunden-analysen gefiltert bzw. verglichen werden (z. B. UK vs. Non-UK).

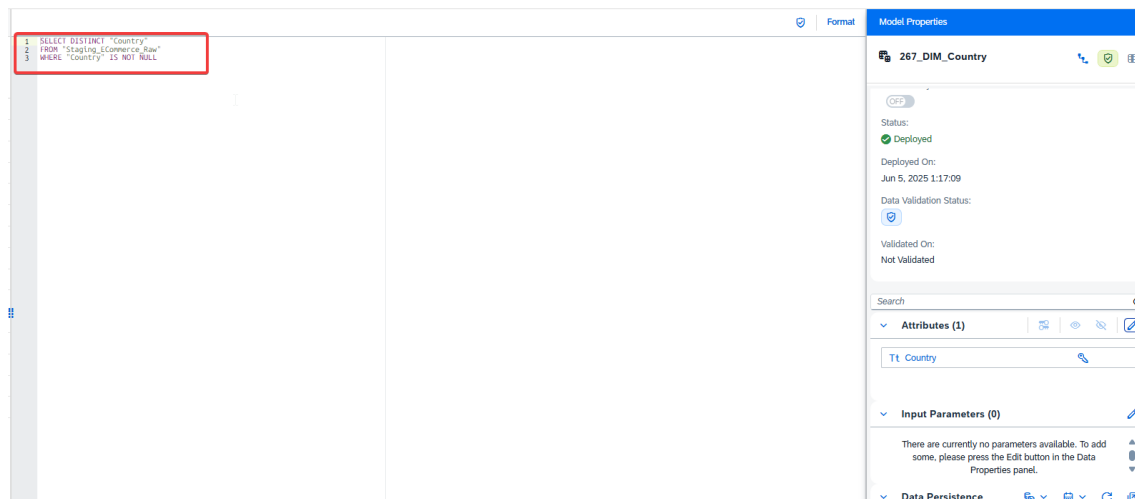
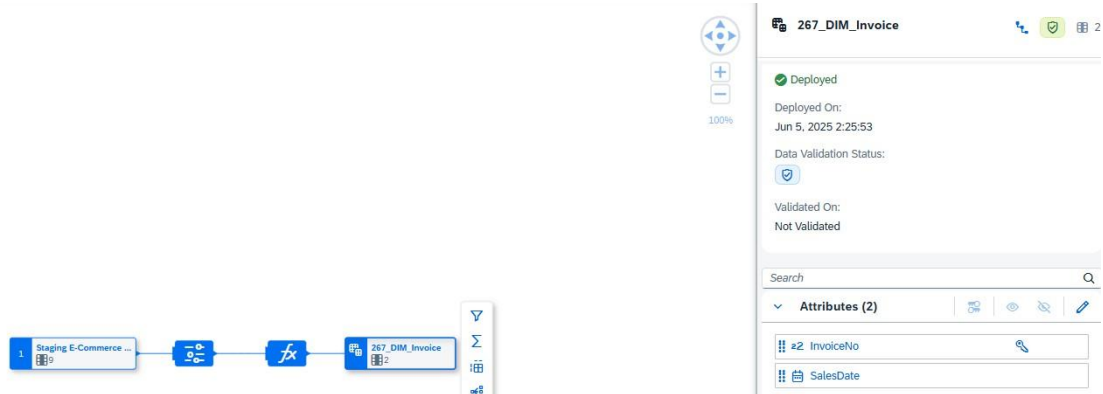


ABBILDUNG 11: CUSTOMER-DIMENSION SQL VIEW¹²



Invoice-Dimension (DIM_Invoice)

1. *Graphical View* erstellen, Key = InvoiceNo.
2. Projection: InvoiceNo, InvoiceDate.
3. Calculated Column SalesDate = TO_DATE(CAST("SalesDate" AS VARCHAR), 'YYYY-MM-DD HH24:MI:SS')
4. Semantic Type:
 - InvoiceNo → *ID*
 - SalesDate → *Date*.

¹² Abdulhadi/Github (2025).

ABBILDUNG 12: INVOICE-DIMENSION¹³

4.4 Fakten-View

Schritt	Aktion	Screenshots/Codesnip
1	Projection auf STG-Tabelle	
2	fx-Column <i>SalesDate</i> → TO_DATE	TO_DATE(CAST("InvoiceDate" AS VARCHAR), 'YYYY-MM-DD HH24:MI:SS')
3	fx-Column <i>TotalPrice</i> → Quantity * UnitPrice	Quantity * UnitPrice
4	Filter: Quantity > 0	

¹³ Abdulhadi/Github (2025).


5	Semantic U age = Fact, Associations zu DIM_PROD- UCT, DIM_CUS- TOMER, DIM_COUN- TRY, DIM_IN- VOICE, TIME_DI- MENSION	
---	---	--

ABBILDUNG 13: ASSOCIATIONS - FACTS

4.5 Kennzahlen & Filter-Variabeln

Im Analytic Model **AM_ECOM_SALES** werden folgende Measures & Filter definiert:

NetSales	TotalPrice	Basisumsatz
ItemsSold	Quantity	Stückzahl
AvgBasket	NetSales / DISTINCT	Ø Warenkorb



Measures (6)
✱ AvgBasket
✱ DistinctInvoices
✱ ItemsSold
✱ NetSales
📦 Quantity
📦 TotalPrice

ABBILDUNG 14: MEASURES - ANALYTIC MODEL¹⁴

¹⁴ Abdulhadi/Github (2025).

Analytic Model Properties
✕

{·} **Filter Variable**

267_AM_ECOM_SALES / Year

Business Name: *

Year

Technical Name: *

YEAR_InvoiceDate

Data Type:

String(4)

Dimension:

Year (SalesDate)
▼

▼ **Value Definition**

Variable filled by:

Manual Input
▼

Filter Type:

Multiple Single Values
▼

Default Value:

2011
✕

Add

▼ **Used in (1)**

▼
Filter on Year (SalesDate)
Filter on Year (SalesDate)
>

▼ **Settings**

☐ Mandatory

ABBILDUNG 15: FILTER - YEAR¹⁵

¹⁵ Abdulhadi/Github (2025).

Insgesamt ergibt sich mit allen Dimensionen, Fakten und entsprechenden Filterungen und weiteren Modifizierungen, folgendes Datenmodell.



ABBILDUNG 16: DATENMODELL¹⁶

4.6 Story-Erstellung in SAP Analytics Cloud

Netto-Umsatz pro Monat (Liniendiagramm)

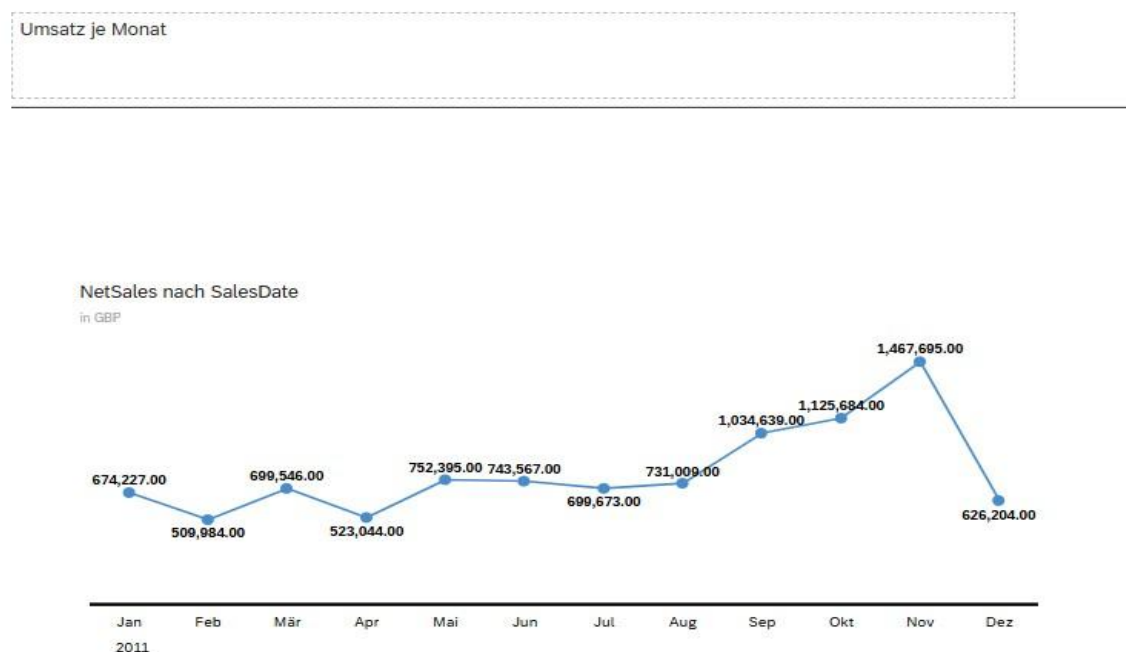


ABBILDUNG 17: NETTOUMSATZ PRO MONAT - LINIENDIAGRAMM¹⁷

Der Jahresverlauf zeigt ein typisches **Saisonmuster** für den Online-Handel:

¹⁶ Abdulhadi/Github (2025).

¹⁷ Abdulhadi/Github (2025).

- Von Januar bis August schwankt der Umsatz moderat zwischen ca. 500 k GBP und 730 k GBP.
- **September–November** steigt die Kurve deutlich an (Peak $\approx 1,47$ Mio GBP) – klassischer Vorweihnachts-/Black-Friday-Effekt.
- Im **Dezember** fällt der Umsatz zwar zurück, bleibt aber höher als im Frühjahr.
⇒ Das Unternehmen erzielt rund **40 % des Jahresumsatzes im letzten Quartal**; Marketing-Budget sollte daher auf Q4 fokussiert werden.

Top 5 Meistverkaufte Produkte (Säulendiagramm)

Top 5 Meistverkaufte Produkte

DistinctInvoices nach CustomerID, InvoiceNo
in GBP

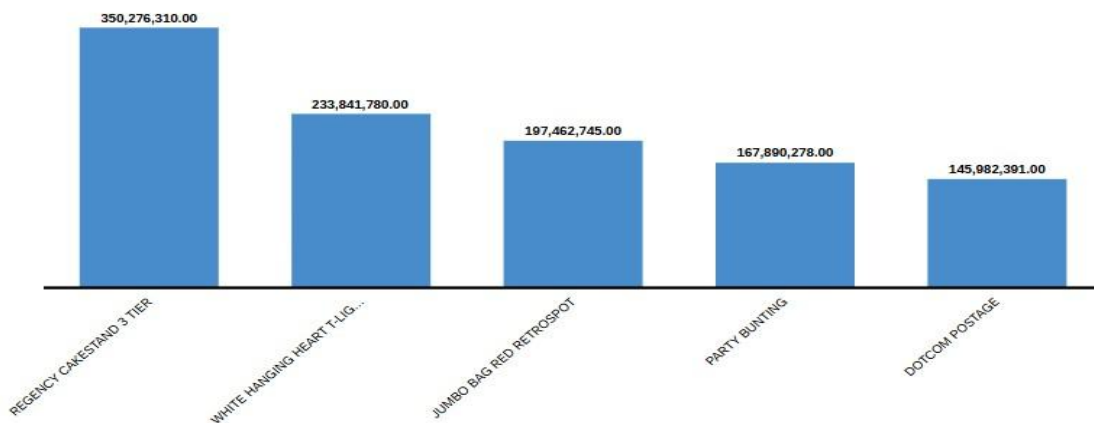


ABBILDUNG 18: TOP 5 MEISTVERKAUFTE PRODUKTE - SÄULENDIAGRAMM¹⁸

Die Balken zeigen, dass das meistverkaufte Produkt einen Umsatz von rund 390 Millionen GBP generiert – deutlich mehr als der zweitplatzierte Artikel mit 233 Millionen GBP. Damit trägt dieses eine Produkt überproportional zum Gesamtumsatz bei.

- Das Spitzenprodukt trägt > **20 %** zum Gesamtumsatz der Top-5 bei – ein klarer Kandidat für Preis- und Lagerbestandsoptimierung.
- Sortimentspflege sollte sich nicht nur auf den Bestseller, sondern auf die **breite Umsatzbasis der Plätze 2-5** konzentrieren, da hier weiteres Potenzial bei geringerer Abhängigkeit besteht.

¹⁸ Abdulhadi/Github (2025).

Durchschnittlicher Einkaufswert nach Land (Säulendiagramm)

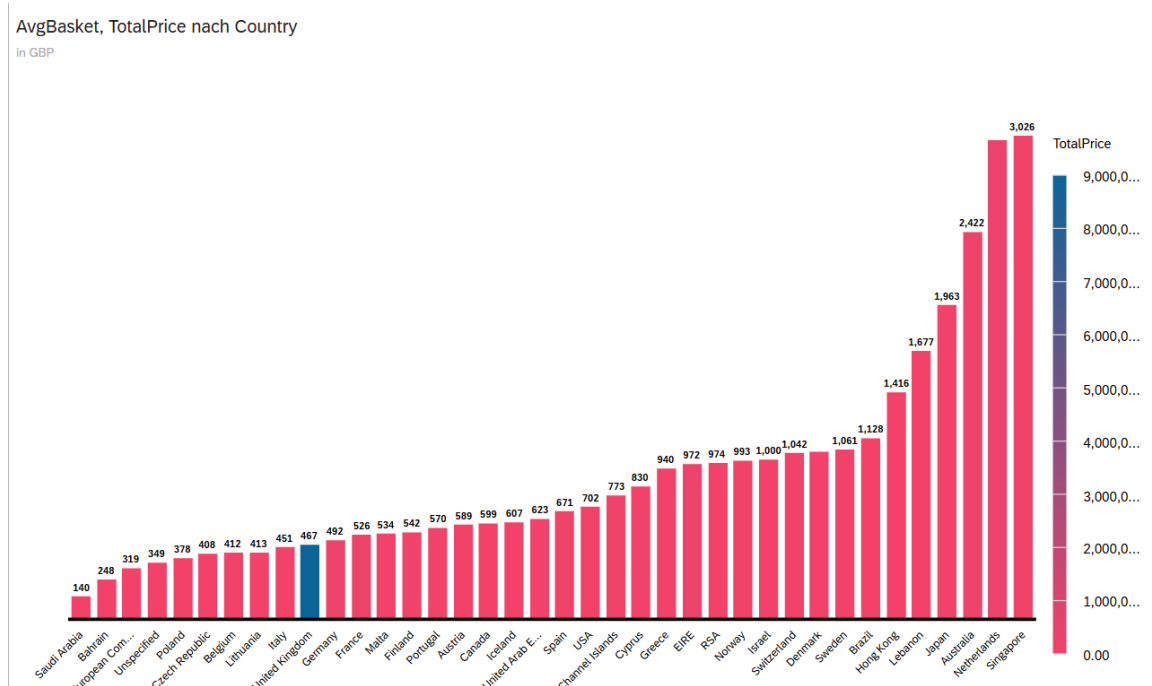


ABBILDUNG 19: DURCHSCHNITTLICHER EINKAUFSWERT NACH LAND - SÄULENDIAGRAMM¹⁹

Hohe Clusterung: Die meisten Länder liegen zwischen 500 – 2 000 GBP Ø-Einkaufswert

- Einzelne Ausreißer (rechte Seite)
- Länder mit blau markiertem Balken (UK-Heimatmarkt) rangieren knapp im Mittelfeld – kein Top-Spender, aber solide Basis.
- ⇒ Für Upselling-Strategien lohnen sich Premium-Ländergruppen (> 1,5 k GBP Ø-Basket) sowie gezielte Kampagnen, um Niedrig-performer in der linken Range anzuheben.

¹⁹ Abdulhadi/Github (2025).

5 Fazit

Die vorliegende Hausarbeit demonstriert, wie sich ein frei zugänglicher E-Commerce-Datensatz mithilfe von **SAP Datasphere** und **SAP Analytics Cloud (SAC)** in einen konsistenten Analyse-Stack überführen lässt – von der Rohdatei bis zum interaktiven Dashboard.

Im ersten Schritt wurde gezeigt, dass sich selbst heterogene CSV-Quellen nach einer verhältnismäßig schlanken Python-Vorbereitung verlustfrei in den Cloud-Space laden lassen. Die **dimensionale Modellierung** (Product, Customer, Country, Invoice, Date) sowie die saubere Trennung von Fakten und Dimensionen erwiesen sich dabei als zentraler Faktor für Performance und Wiederverwendbarkeit. Durch eindeutige **Semantic-Typen** und **Text-Associations** konnten IDs automatisch mit Klartext ergänzt werden, was der späteren Berichtsqualität unmittelbar zugutekommt.

Das auf dem Fact-View basierende **Analytic Model** stellte alle notwendigen Kennzahlen – Gesamtumsatz, verkaufte Stückzahl, durchschnittlicher Warenkorb etc. – in wiederverwendbarer Form bereit. Besonders wertvoll waren dabei Variablen wie *Year*, die eine dynamische Filterung ohne zusätzliche Modellversionen ermöglichen.

Die anschließende Auswertungsphase in der SAC bestätigte die Güte des Datenmodells:

- Das Liniendiagramm zum **monatlichen Umsatz** offenbarte ein klares Saisonmuster mit einem massiven Peak im vierten Quartal, was datengetriebenes Budget-Shifting rechtfertigt.
- Die **Top-5-Produktanalyse** zeigte eine signifikante Abhängigkeit von einem einzelnen Artikel – ein konkreter Ansatzpunkt für Sortiments- und Risikomanagement.
- Die **Länderanalyse** belegte stark divergierende Warenkorbwerte und identifizierte Premium-Cluster, die sich gezielt für Upselling-Kampagnen eignen.

In Summe belegt das Projekt, dass mittelständische Online-Händler durch den Einsatz von SAP-Cloud-BI-Werkzeugen in kurzer Zeit ein belastbares, skalierbares Reporting aufbauen können – ohne tiefgreifende ABAP- oder BW-Kenntnisse. Die End-to-End-Cloud-Pipeline reduziert manuelle Aufwände, erhöht die Datenqualität und liefert Entscheidungen nahezu in Echtzeit geschäftsrelevante Insights.

Literaturverzeichnis

Abdulhadi, K./*Github* (2025): nichtkarim/BI2_Projekt, in:

https://github.com/nichtkarim/BI2_Projekt, [Stand 04.06.25].

Carrie (2017): E-Commerce Data, in: <https://www.kaggle.com/datasets/carrie1/ecommerce-data/data>, [Stand 04.06.25].

Corporation, A. (2024): Is the On-Premises Data Warehouse Dead?, in: <https://www.action.com/blog/data-warehouse/is-the-on-premises-data-warehouse-dead/>, [Stand 05.06.25].

Kimball, R./Ross, M. (2002): The data warehouse toolkit. The complete guide to dimensional modeling, 2. Aufl., New York, NY, Weinheim.