

Machine Learning mit Python

Ziel: Die Integration von Machine Learning in Qlik Sense unter Verwendung von Python und den Bibliotheken Pandas, NumPy, Matplotlib, Seaborn und Scikit-Learn bietet zahlreiche Vorteile.

Hier sind einige der wichtigsten Vorteile im Detail:

1. Datenmanipulation und -vorbereitung:

Pandas und NumPy bieten leistungsstarke Werkzeuge zur Datenmanipulation, -bereinigung und -vorbereitung, die für Machine Learning unerlässlich sind.

2. Effiziente Berechnungen:

NumPy ermöglicht effiziente numerische Berechnungen und Verarbeitung großer Datenmengen.

3. Umfangreiche Visualisierungsmöglichkeiten:

Mit Matplotlib und Seaborn können Sie komplexe Datenvisualisierungen erstellen, die Ihnen helfen, Muster und Trends in den Daten zu erkennen.

4. Einfache Integration:

Diese Bibliotheken lassen sich leicht in Python-Skripte integrieren und bieten eine Vielzahl von Diagrammtypen und Anpassungsmöglichkeiten.

5. Umfassende Bibliothek:

Scikit-Learn bietet eine breite Palette von Algorithmen für Klassifikation, Regression, Clustering und Dimensionalitätsreduktion.

6. Einfachheit und Konsistenz:

Die einheitliche API von Scikit-Learn macht es einfach, verschiedene Machine Learning-Algorithmen auszuprobieren und zu vergleichen.

7. Werkzeuge zur Modellbewertung:

Scikit-Learn bietet zahlreiche Werkzeuge zur Modellbewertung und -optimierung, wie Kreuzvalidierung und Grid-Suche.

8. CSV- und Excel-Unterstützung:

Sie können Daten problemlos zwischen Qlik Sense und Python austauschen, indem Sie CSV- oder Excel-Dateien exportieren und importieren.

9. Qlik Sense APIs:

Die Verwendung der Qlik Sense APIs ermöglicht eine noch nahtlosere Integration und Automatisierung des Datenflusses zwischen den Plattformen.

10. Automatisierte Workflows:

Python-Skripte können automatisiert werden, um regelmäßig Datenanalysen und Modelltrainings durchzuführen, was Zeit spart und die Reproduzierbarkeit erhöht.

11. Wiederverwendbare Skripte:

Einmal erstellte Skripte können leicht wiederverwendet und angepasst werden, um neue Anforderungen zu erfüllen.

12. Erweiterte Vorhersagen:

Durch die Integration von Machine Learning-Modellen können Sie fortschrittliche Vorhersagen und Prognosen direkt in Ihren Qlik Sense-Dashboards anzeigen.

13. Interaktive Visualisierungen:

Kombinieren Sie die interaktiven Visualisierungsmöglichkeiten von Qlik Sense mit den analytischen Fähigkeiten von Python, um tiefere Einblicke in Ihre Daten zu gewinnen.

Seite 1: Auf der ersten Seite erhält der Benutzer eine Gesamtübersicht über die Anwendung.



Seite 2: Auf dieser Seite kann der Benutzer Daten planen und seine vorgenommenen Änderungen überprüfen.

Planung

Lade Vorlage-Daten (Plan)

Planen (Genauigkeit des ML-Modells: 83,35%)

City (Plan)	Product (Plan)	Price/pc. (Plan)	Personell (Plan)	Number of Machines (Plan)	R&D (Plan)	Customer Support (Plan)	Marketing & Advertising (Plan)	Sales Quantity (Plan)
Frankfurt	Banduhr	0,45	2350000	550	790000	370000	1030000	38663620
Hamburg	Banduhr	0,75	2150000	625	550000	370000	590000	32808228
Stuttgart	Banduhr	0,75	2750000	500	750000	370000	860000	33616017
Berlin	Banduhr	0,78	2350000	530	500000	300000	690000	30596082
Berlin	E-Bike	0,78	2850000	525	500000	400000	550000	34099258
Stuttgart	Roller	0,89	1560000	625	720000	450000	550000	26282672

1 - 20 of 20

Ändern

Anz. Training-Datensätze

3.000

Fehlerquote des ML-Modells

16,65%

Anz. Änderungen

50

Planung

Kontrolle

Änderungen Zurücksetzen

Python ML-Modell Ausführen

Kontrolle

Lade Vorlage-Daten (Plan)

...	Pr...	Preis pro Stück (Actual)	Preis pro Stück (Plan)	Personal (Actual)	Personal (Plan)	Anz. Maschinen (Actual)	Anz. Maschinen (Plan)	R&D (Actual)	R&D (Plan)	Kundensup... (Actual)	Kundensup... (Plan)	Marketing & Werbung (Actual)	Marketing & Werbung (Plan)
tals				59580000	517200...	10365	10530	12990...	1290...	7080000	7280000	15430000	16080000
rlin	Radio	0,21	0,25	2610000	2650000	620	620	430000	500000	410000	400000	1030000	950000
inkf...	E-Bike	0,42	0,35	2120000	2000000	520	450	790000	800000	330000	350000	510000	800000
inkf...	Band...	0,60	0,45	2350000	2350000	425	550	790000	790000	370000	370000	1030000	1030000
inkf...	Roller	1,34	1,34	1810000	1900000	420	420	790000	750000	330000	330000	1030000	1030000
inkf...	Radio	0,15	0,15	2100000	2150000	620	620	720000	720000	330000	330000	860000	860000
inkf...	Halsb...	0,31	0,31	3670000	3500000	425	425	650000	650000	330000	330000	690000	690000
imb...	Roller	1,42	1,55	1820000	1820000	530	530	650000	650000	370000	370000	1030000	1030000
imb...	Band...	0,75	0,75	2360000	2150000	625	625	650000	550000	370000	370000	510000	590000
imb...	Halsb...	0,40	0,25	4200000	4200000	525	450	570000	600000	330000	350000	860000	860000
imb...	Radio	0,10	0,20	2370000	2400000	515	500	570000	450000	370000	350000	860000	1050000
imb...	E-Bike	0,88	0,95	2350000	2350000	630	630	500000	550000	290000	290000	510000	550000
ittg...	E-Bike	1,09	1,25	2620000	2620000	520	520	790000	790000	410000	410000	1030000	950000
ittg...	Halsb...	0,20	0,10	3950000	3950000	420	420	790000	790000	370000	370000	510000	510000
ittg...	Radio	0,17	0,17	2600000	2900000	430	550	720000	750000	330000	350000	860000	1000000

Anz. Training-Datensätze

3.000

Fehlerquote des ML-Modells

16,65%

Anz. Änderungen

50

Planung

Kontrolle

Änderungen Zurücksetzen

Python ML-Modell Ausführen

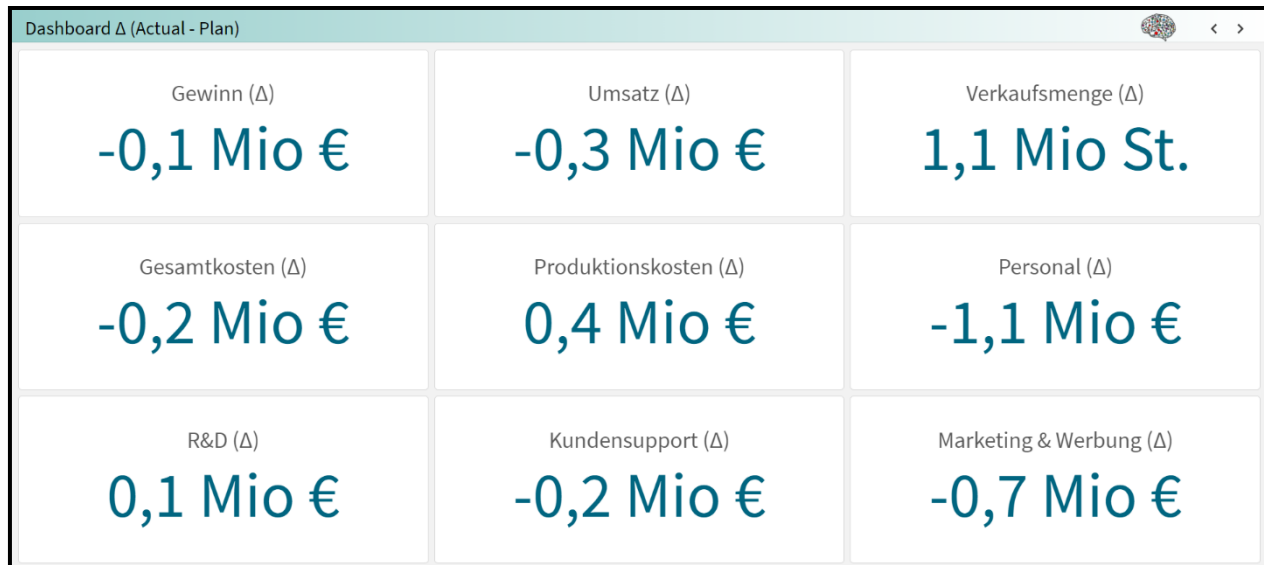
Seite 3: Auf dieser Seite werden die aktuellen Daten visualisiert.

Dashboard (Actual)		
Gewinn (Actual)	Umsatz (Actual)	Verkaufsmenge (Actual)
280,3 Mio €	440,0 Mio €	728,3 Mio St.
Gesamtkosten (Actual)	Produktionskosten (Actual)	Personal (Actual)
159,8 Mio €	104,5 Mio €	50,6 Mio €
R&D (Actual)	Kundensupport (Actual)	Marketing & Werbung (Actual)
13,0 Mio €	7,1 Mio €	15,4 Mio €

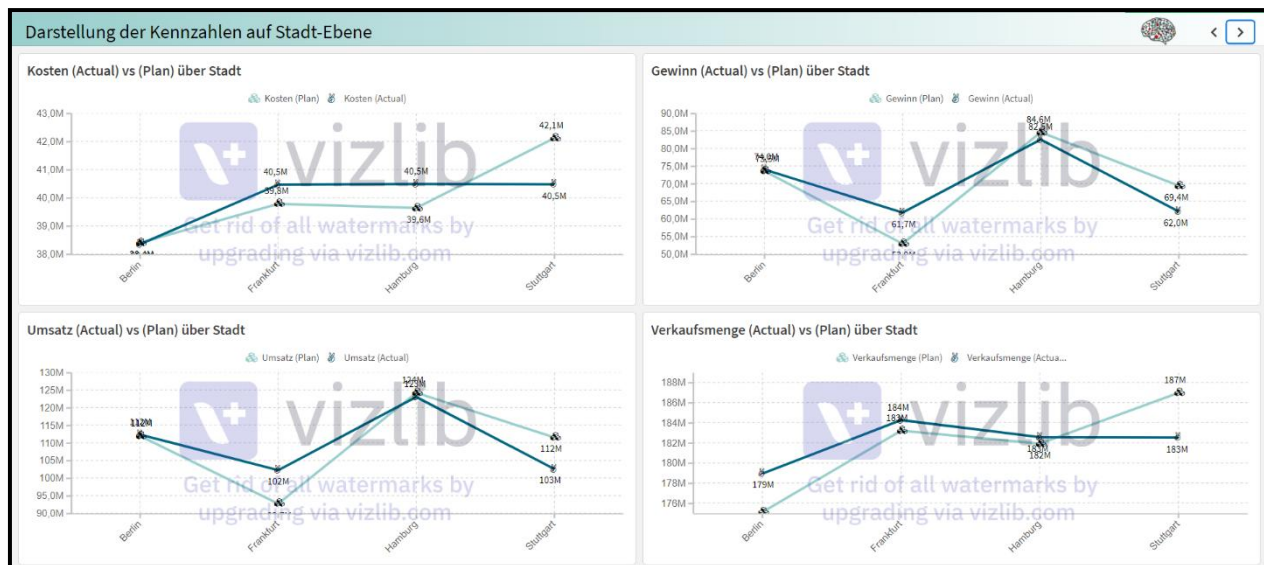
Seite 4: Auf dieser Seite werden die Plandaten visualisiert.

Dashboard (Plan)		
Gewinn (Plan)	Umsatz (Plan)	Verkaufsmenge (Plan)
280,4 Mio €	440,3 Mio €	727,2 Mio St.
Gesamtkosten (Plan)	Produktionskosten (Plan)	Personal (Plan)
159,9 Mio €	104,1 Mio €	51,7 Mio €
R&D (Plan)	Kundensupport (Plan)	Marketing & Werbung (Plan)
12,9 Mio €	7,3 Mio €	16,1 Mio €

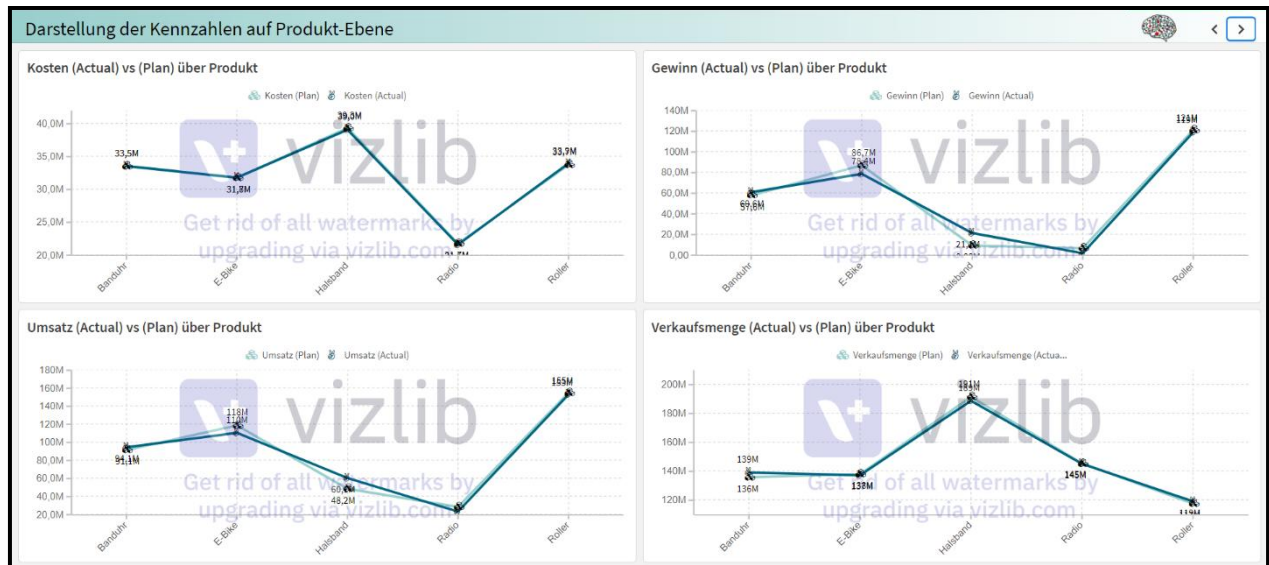
Seite 5: Auf dieser Seite wird die Differenz zwischen Ist-Daten und Plandaten visualisiert.



Seite 6: Auf dieser Seite werden die Ist-Daten und Plandaten auf geografischer Ebene visualisiert.



Seite 7: Auf dieser Seite werden die Ist-Daten und Plandaten auf Produktebene visualisiert.



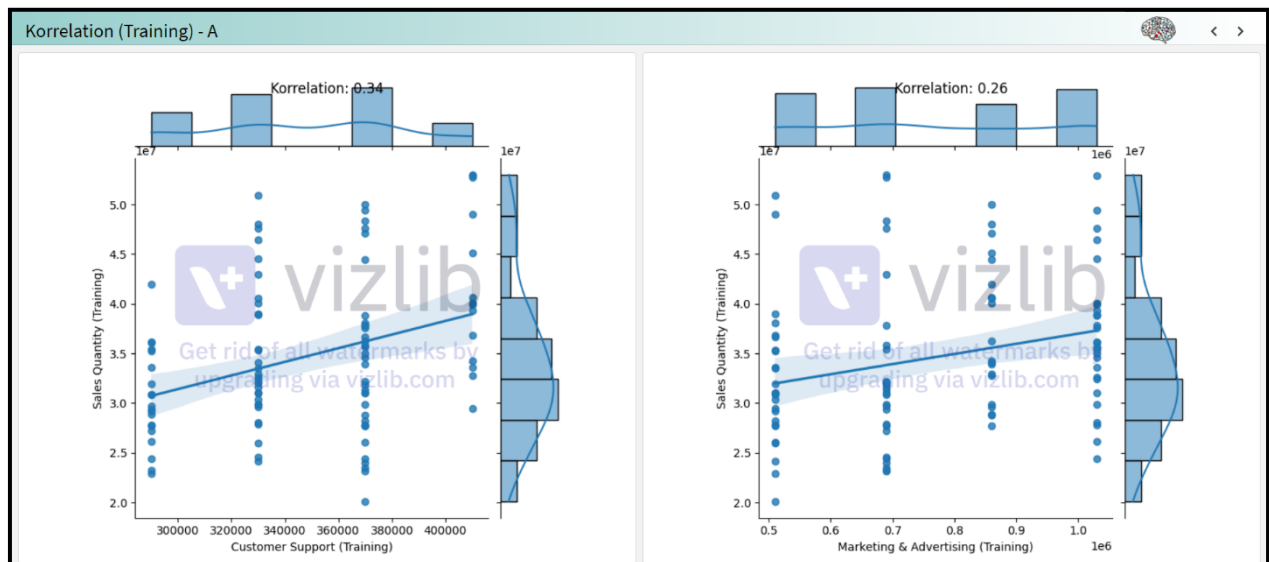
Seite 8: Auf dieser Seite werden die Korrelationen zwischen numerischen Kennzahlen grafisch dargestellt.



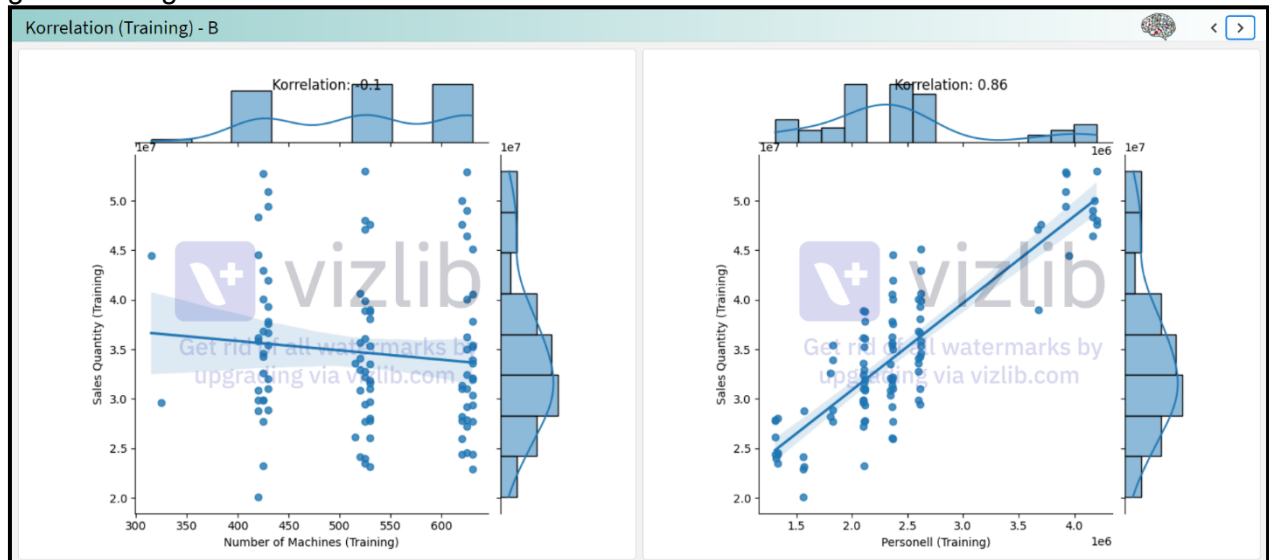
Seite 9: Auf dieser Seite werden wichtige Kennzahlen zu den numerischen Daten grafisch dargestellt.

Statistiken								
Statistiken (Actual)								
	count	max	mean	min	std	25%	50%	75%
Customer Support (Actual)	20	410000	354000	290000	30157,481397002	330000	370000	370000
Marketing & Advertising (Actual)	20	1030000	771500	510000	221484,82186324	510000	860000	1030000
Number of Machines (Actual)	20	630	518,25	420	83,197213389569	425	520	620
Personell (Actual)	20	4200000	2529000	1560000	768744,020088	2100000	2355000	2620000
Price/pc. (Actual)	20	1,43	0,6435	0,1	0,43099975576987	0,285	0,6	0,8825
R&D (Actual)	20	790000	649500	430000	111943,35973819	570000	650000	737500
Sales Quantity (Actual)	20	51007084	36412856,7	24184636	6911810,9009621	33708695,75	35158150	38863373
Statistiken (Plan)								
	count	max	mean	min	std	25%	50%	75%
Customer Support (Plan)	20	410000	354000	290000	30157,481397002	330000	370000	370000
Marketing & Advertising (Plan)	20	1030000	771500	510000	221484,82186324	510000	860000	1030000
Number of Machines (Plan)	20	720	523,25	420	92,156545873684	425	520	621,25
Personell (Plan)	20	4200000	2529000	1560000	768744,020088	2100000	2355000	2620000
Price/pc. (Plan)	20	1,43	0,6485	0,15	0,42489967856202	0,285	0,6	0,8825
R&D (Plan)	20	790000	649500	430000	111943,35973819	570000	650000	737500
Sales Quantity (Plan)	20	49961461	36410985,85	24936714	7004193,3028681	32626652,75	34127447	39602340,25

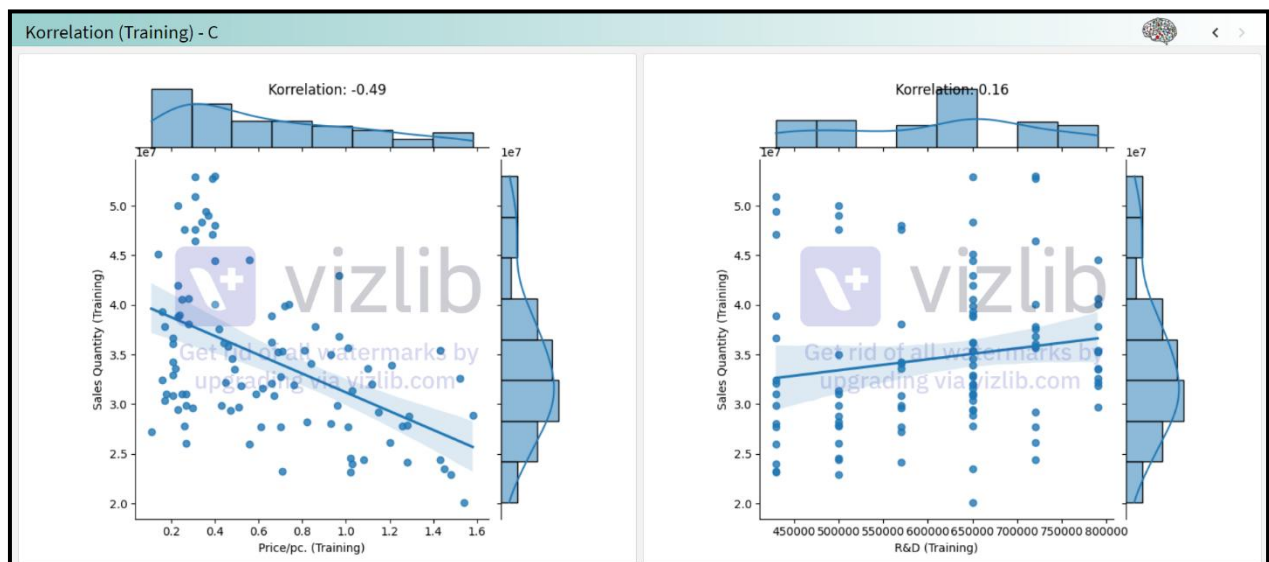
Seite 10: Auf dieser Seite wird die Beziehung zwischen Kennzahlen in Form eines Pairplots grafisch dargestellt.



Seite 11: Auf dieser Seite wird die Beziehung zwischen Kennzahlen in Form eines Pairplots grafisch dargestellt.



Seite 12: Auf dieser Seite wird die Beziehung zwischen Kennzahlen in Form eines Pairplots grafisch dargestellt.



Datenmodell: Dieser Ausschnitt repräsentiert das Datenmodell, in dem Tabellen miteinander verknüpft sind.

