



# MUSTERBASIERTE KEYWORD-RECHERCHEN MIT KNIME FÜR ANFÄNGER UND FORTGESCHRITTENE

Rebecca Schwarz

Eine alltägliche Aufgabe in der SEO: die Keyword-Recherche. Das allgemeine Ziel: Erkenntnisse darüber zu gewinnen, welche Suchphrasen von Nutzern über Suchmaschinen gesucht werden und in welche Aspekte sich Themen gliedern. Auch die Wortwahl der Zielgruppe ist hierbei spannend: Wird für den Kauf von Produkten beispielsweise eher mit dem Verb „kaufen“ oder dem Verb „bestellen“ gesucht? Damit die Antwort nicht immer „It depends“ lautet, können diese Abwägungen für gleichförmige Inhalte über musterbasierte Recherchen analysiert werden. Dazu werden verschiedene Suchkombinationen geprüft und schlussendlich wird verglichen, welche Kombinationen auf Basis des Suchvolumens am relevantesten sind. Wie sich diese Analyse leicht umsetzen und mit unendlich vielen Kombinationen von Suchanfragen wiederholen lässt, zeigt der nachfolgende Artikel von Rebecca Schwarz mit dem kostenlosen Datenanalyse-Tool KNIME.

## DIE AUTORIN



Rebecca Schwarz ist SEO-Consultant bei der get:raction GmbH. Ihr Arbeitsalltag dreht sich um die Konzeption von SEO-Strategien und die Unterstützung von Kunden in der redaktionellen SEO. Um größere Datens Mengen effizient zu verarbeiten und um bei wiederkehrenden SEO-Tasks Zeit zu sparen, nutzt sie die Open-Source-Software KNIME.

In der folgenden Schritt-für-Schritt-Anleitung wird gezeigt, wie sich Keywords miteinander kombinieren lassen. Als Beispiel werden verschiedene Produktbenennungen mit transaktionalen Verben kombiniert. Dann wird das Suchvolumen der neuen Kombinationen abgefragt. Hierzu wird gezeigt:

1. wie die Erstellung von CSV-Dateien mit allen Suchkombinationen funktioniert, um das Suchvolumen im Google-Ads-Keyword-Planer

via Upload-Funktion hochzuladen und abzufragen.

2. wie die automatische Abfrage des Suchvolumens innerhalb von KNIME mithilfe der SISTRIX-API realisiert werden kann.

Der Workflow endet jeweils mit einer Gruppierung der Daten, um die gefragten Kombinationen nach Produkt und nach transaktionalen Verb zu erhalten, wie es aus Pivot-Tabellen in Excel bekannt ist.

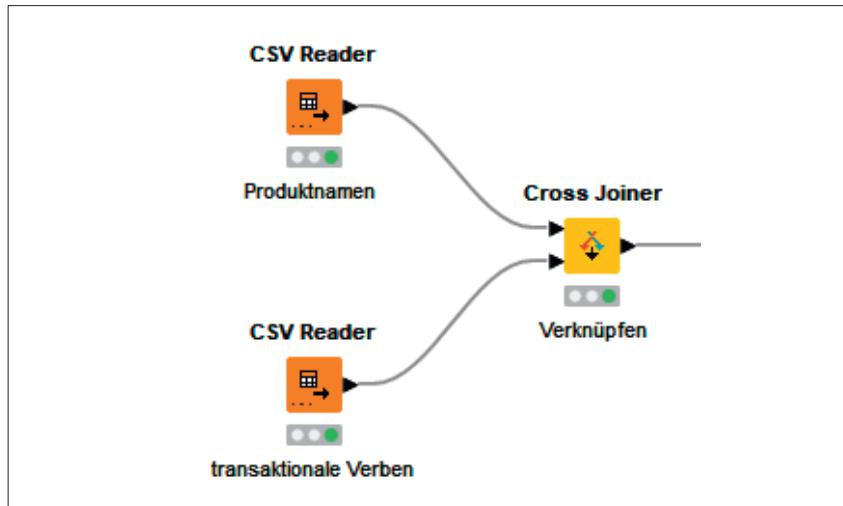


Abb. 1: Verbinden der Knoten CSV READER und CROSS JOINER

## Hinweis vor dem Start in den Workflow

Das hier vorgestellte Beispiel bezieht sich auf transaktionale Suchanfragen, bestehend aus einem Produktnamen und einem transaktionalen Verb. Der Workflow ist jedoch überall anwendbar, wo gleichförmige Suchanfragen eines Themengebiets analysiert werden sollen. Beispiele sind:

- » [Art von Geschäft] + [Stadtname] wie „friseur frankfurt“
- » [Ländername] + [Thema] wie „brasilien schulsystem“
- » [Farbe] + [Produktname] wie „blaue schuhe“

## Start der Analyse in KNIME

Für den Start der Analyse werden die zu kombinierenden Anfragen benötigt. Dazu werden die Produktnamen

und die transaktionalen Verben jeweils in eine eigene CSV-Datei geschrieben. Zu wenige Ideen für die Findung der Verben und zu wenig Zeit für die Erstellung der CSV-Datei? Einfach kurz ChatGPT darum bitten!

Sind die CSV-Dateien erstellt, kann die Arbeit in KNIME auch direkt starten. Dazu werden die CSV-Dateien einfach auf die grafische Oberfläche in KNIME gezogen. Automatisch wird dadurch jeweils ein Knoten, der CSV READER, erstellt. Mit Doppelklick auf die Knoten öffnet sich ein Dialogfenster mit einer Vorschau der Daten. Werden in der Vorschau die Umlaute falsch angezeigt, sollte im Reiter „Encoding“ die Codierung auf UTF-8 oder UTF-16 umgestellt werden. Mit Klick auf „OK“ und einem anschließenden Rechtsklick auf den jeweiligen Knoten und der Auswahl „EXECUTE“ ändert sich unterhalb des Knotens die Ampelanzeige von Gelb zu Grün. Sind beide CSV READER mit „EXECUTE“ ausgeführt, können die Daten schon kombiniert werden.

### TIPP

#### Hilfe für Anfänger:

Alle Knoten in KNIME haben an ihren Enden Ports, aus denen Daten hinein- und herausfließen können. Knoten, die Daten einlesen, haben nur Ausgangsports. Knoten, die Daten wegschreiben und außerhalb der Oberfläche zur Verfügung stellen, haben nur Eingangssports. Damit Daten problemlos in KNIME weitergegeben werden, müssen die entsprechenden Ports immer über die Pfeilsymbole miteinander verbunden werden.

## Kombination der Daten mit dem CROSS JOINER

Um die Daten der beiden Knoten miteinander zu kombinieren, wird der sogenannte CROSS JOINER benötigt. Dieser führt, wie der Name schon sagt, einen Cross Join aus. Das bedeutet, dass die Daten der ersten Datei mit allen Daten der zweiten Datei kom-

## WARUM WIRD KNIME VERWENDET?

KNIME bietet als kostenfreie Open-Source-Software die Möglichkeit, eine Vielzahl an Datenanalysen durchzuführen. Der Vorteil für SEO-Experten: Es sind keine Programmierkenntnisse notwendig und einmal gebaute Workflows können ganz einfach mit neuen Daten wiederholt, angepasst oder auch an Kollegen zur Bearbeitung weitergegeben werden. Die sogenannten Workflows bestehen aus verschiedenen Knoten, die einzelne Aufgaben im Datenprozess darstellen. Durch die aneinandergereihten Knoten können einmal ins Tool geladene Daten durchfließen und entsprechend dem jeweiligen Knoten bearbeitet werden. Wichtig zu erwähnen: Die Software ist in jedem Fall für Anfänger geeignet und ein genauerer Blick in die vielen Funktionen lohnt sich, um im Arbeitsalltag Zeit und Nerven zu sparen.

In der Website Boosting veröffentlichen wir bereits seit Ausgabe 53 immer wieder hilfreiche Problemlösungen mit KNIME. 18 davon sind online frei verfügbar und downloadbar. Suchen Sie einfach in der Rubrik Artikel mit dem Stichwort „KNIME“ danach oder nutzen Sie unsere vorkonfigurierte Such-URL: [einfach.st/knimeartikel](http://einfach.st/knimeartikel).

**Do it: Eine einfache Analyse von Search-Console-Daten mit KNIME**  
Datenreise durch KNIME und alle Umformungen plus die Zusammensetzung geht ab jetzt vollautomatisch und in Echtzeit. Bereits in Ausgabe 53 wurde das Thema KNIME aufgegriffen...

RESSORT: Content & Systeme AUTOR: Michael Höherleiter JAHR: 2021 AUSGABE: 56

**Beyond Crawling: Individuelle Klassifizierung Ihrer Webseiten nach Typ**  
der Roboter eines Crawls und KNIME schnell und einfach nachgeholt werden.

Screaming-Frog-Daten in KNIME importieren. Die Grundlagen von KNIME wurden bereits in den...

RESSORT: Web Controlling AUTOR: Michael Höherleiter JAHR: 2021 AUSGABE: 56

**Datenabfrage-Abfrage leicht gemacht: mit KNIME und der Sistrix API**  
Execute werden die Daten in KNIME importiert. Um schnell sehen zu können, wie die Daten in KNIME aussehen, aktivieren Sie im Menü von KNIME unter dem Punkt View den...

RESSORT: Web Controlling AUTOR: Michael Höherleiter JAHR: 2022 AUSGABE: 77

**Data Wrangling mit KNIME: Screaming-Frog-Crawls für Analysen aufbereiten**

wesentlichen Vorteil von KNIME kann man anhand des Screenshots nicht erkennen, nämlich dass KNIME ohne Probleme mit sehr großen Datennummern zureckkommt. Da KNIME ursprünglich...

RESSORT: Content & Systeme AUTOR: Patrick Lüwer JAHR: 2018 AUSGABE: 53

**Beyond Crawling: Insights durch N-Gramm-Analysen (Teil 4)**

bei Bedarf einfach von KNIME Hub (URL im Beitrag) per Drag and Drop in die Arbeitsfläche von KNIME ziehen. Außerdem bietet KNIME die Möglichkeit...

RESSORT: Content & Systeme AUTOR: Michael Höherleiter JAHR: 2021 AUSGABE: 56

biniert werden. Stehen in Datei eins „haus“ und „auto“ und in Datei zwei die Verben „kaufen“ und „verkaufen“, ergeben sich dadurch vier Kombinationen: „haus kaufen“, „haus verkaufen“, „auto kaufen“ und „auto verkaufen“. Durch einen Cross Join werden also entsprechend mehr Anfragen generiert, wodurch die Anzahl der Kombinationen auch schnell sehr groß werden kann.

Zum Kombinieren wird nur der CSV READER mit den Produktnamen am oberen Port und der CSV READER mit den transaktionalen Verben am unteren Port mit dem CROSS JOINER verbunden (Abbildung 1).

Im CROSS JOINER müssen keine Einstellungen vorgenommen werden. Es genügt, den Knoten mit Rechtsklick auf „EXECUTE“ auszuführen. Erscheint unterhalb des Knotens das grüne Symbol, kann mit Rechtsklick auf „Joined Table“ eine Vorschau der Daten angezeigt werden. Nun sollte es zwei Spalten geben, die alle Kombinationen der Anfragen auflisten.

## Manipulation mit der STRING MANIPULATION

Da die Anfragen nun aber noch in zwei Spalten stehen, müssen die Spalten nun so manipuliert werden, dass die Inhalte in einer Spalte zusammengeführt werden. Hierbei hilft der Knoten STRING MANIPULATION. Dieser Knoten ist sehr mächtig und in der Lage, unzählige Manipulationen an Daten durchzuführen, das Schweizer Taschenmesser der Datenmanipulation. Zunächst wird der Knoten mit dem CROSS JOINER verbunden. Mit Doppelklick auf den Knoten öffnet sich nun das Dialogfenster. In Abbildung 2, Ziffer 1 sind alle Spalten der aktuellen Daten zu sehen. In Abbildung 2, Ziffer 2 sind alle Funktionen gelistet, die im Fenster rechts daneben genau erklärt werden. Unterhalb ist das Feld für die Expression zu sehen (Abbildung 2, Ziffer 3), in dem die Spaltennamen mit den entsprechenden

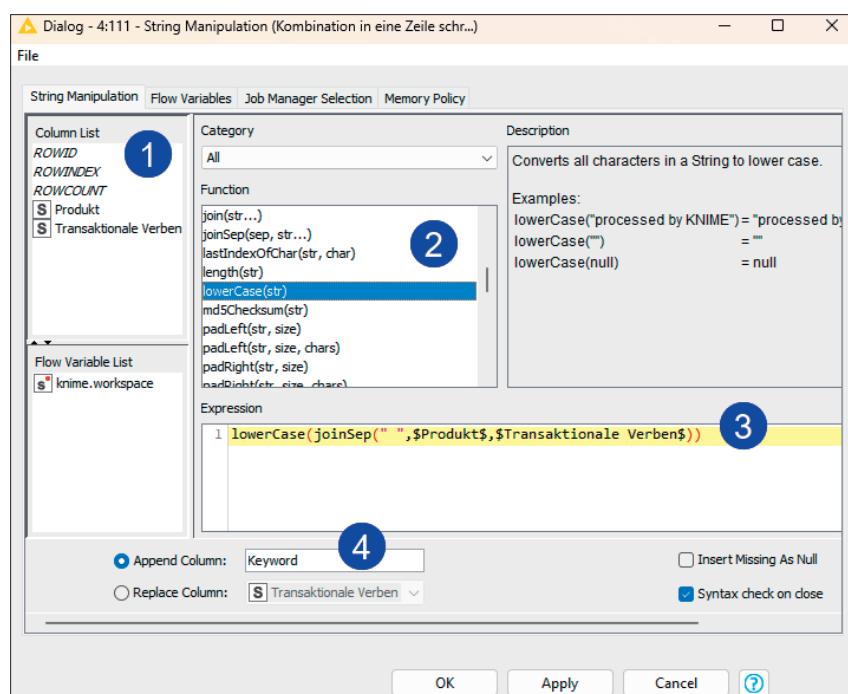


Abb. 2: Konfiguration des Knotens STRING MANIPULATION

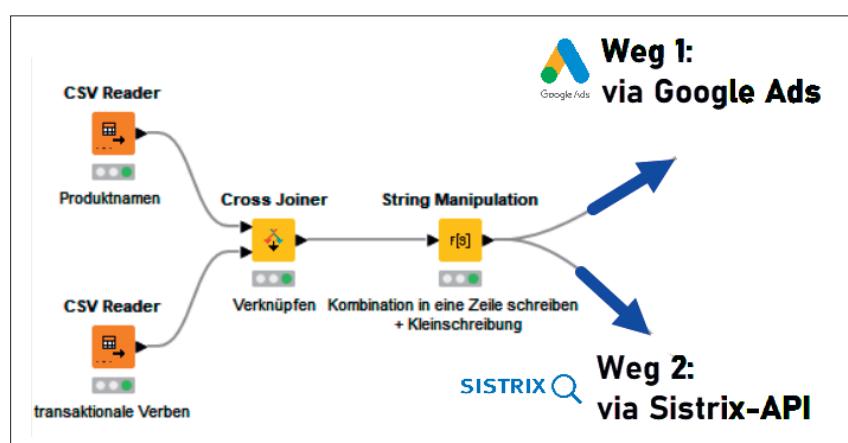


Abb. 3: Verschiedene Wege zum weiteren Vorgehen der Abfrage

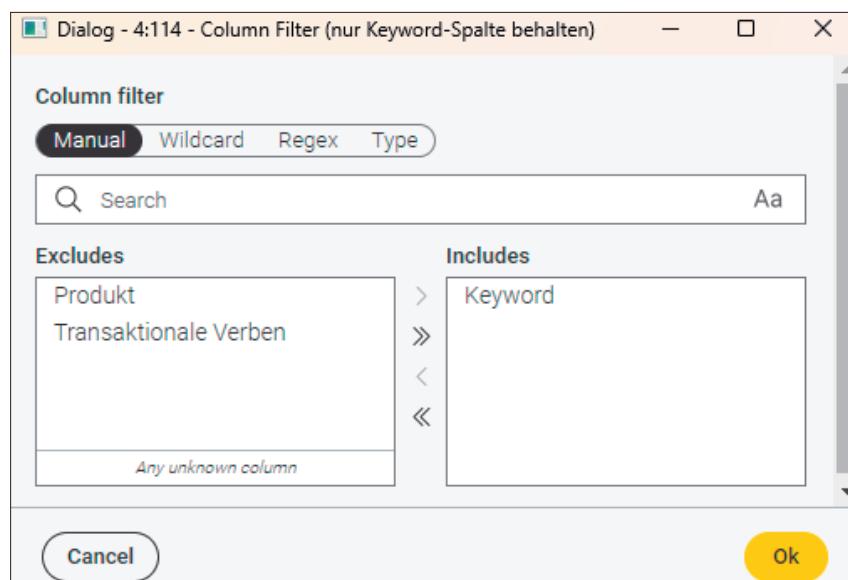


Abb. 4: Keyword-Spalte im Knoten COLUMN FILTER auswählen

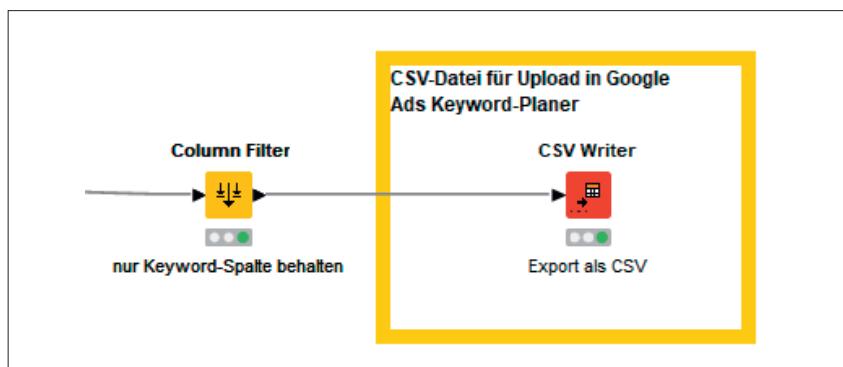


Abb. 5: CSV WRITER zur Erstellung der Datei anhängen

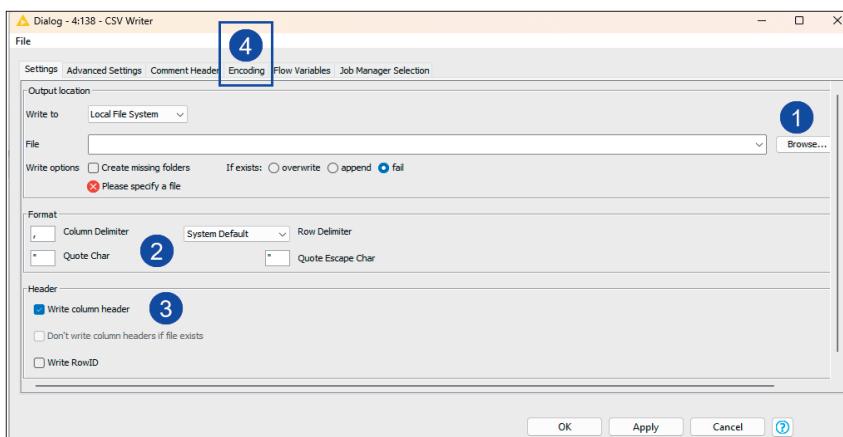


Abb. 6: Konfiguration des Knotens CSV WRITER

**TIPP****Alles easy:**

Der komplette und bereits fertige Workflow ist frei verfügbar online unter [einfach.st/knime86](http://einfach.st/knime86) abrufbar und kann bequem per Drag-and-drop in die eigene KNIME-Umgebung integriert werden. Sie können ihn also selbst Schritt für Schritt aufbauen oder den fertigen Workflow ansehen, damit herumspielen und nach Ihren Wünschen und Parametern gestalten.

Funktionen konfiguriert werden.

Im hier vorliegenden Fall sollen die Inhalte der beiden Spalten mit einem Leerzeichen getrennt in einer neuen Spalte zusammengefasst werden. Für die Abfragen sollen zusätzlich alle Inhalte nur noch in Kleinschreibung vorliegen. Deshalb werden zwei Funktionen kombiniert:

1. **joinSep:** In dieser Funktion müssen erst ein Separator, hier ein Leerzeichen, und die entsprechenden

Spalten zur Kombination ausgewählt werden. Die Funktion lautet deshalb: (joinSep(„ „,\$Produkt\$,\$Transaktionale Verben\$)).

2. **lowerCase:** Mit dieser Funktion wird nun noch der Befehl gegeben, alle Buchstaben in Kleinschreibung zu ändern. Damit dafür keine weitere Zeile notwendig ist, wird diese Funktion um die andere Funktion gelegt. Die finale Expression lautet deshalb: lowerCase(joinSep(„ „,\$Produkt\$,\$Transaktionale Verben\$)).

Im letzten Schritt wird nun noch festgelegt, ob die Manipulation eine bestehende Zeile überschreiben soll (= Replace Column) oder eine neue Zeile ans Datenset angehängt werden soll (= Append Column).

Da im nächsten Schritt eine CSV-Datei für den Upload in Google Ads erstellt werden soll und Google genaue Vorstellungen hat, wie diese aufgebaut sein soll, wird Append Column mit dem Namen „Keyword“ gewählt (Abbildung 2, Ziffer 4).

Ab hier ergeben sich nun die zwei angekündigten Möglichkeiten der Abfrage des Suchvolumens der Keyword-Kombinationen. Weg eins mit der Erstellung von CSV-Dateien und der Abfrage via Google-Ads-Keyword-Planer und Weg zwei mit der Abfrage via SISTRIX-API innerhalb von KNIME (Abbildung 3).

## **Weg eins: Erhebung des Suchvolumens mit CSV-Dateien via Google Ads**

### **Die Erstellung der Google-Ads-Datei**

Um nun die gewünschte CSV-Datei für Google Ads zu erhalten, müssen die nicht mehr benötigten Datenspalten aus dem Datenset gefiltert werden. Dazu wird an die STRING MANIPULATION der Knoten COLUMN FILTER gehängt. Hierbei wird nun in die Includes nur die Spalte Keyword eingefügt (Abbildung 4). Nach der Ausführung des Knotens wird der CSV WRITER angehängt (Abbildung 5).

Im CSV WRITER muss zunächst über „Browse“ gewählt werden, wo die zu erstellende CSV-Datei abgelegt werden soll (Abbildung 6, Ziffer 1). Bei den Format-Einstellungen müssen in den Feldern Quote Char und Quote Escape Char die Werte gelöscht werden (Abbildung 6, Ziffer 2), da ansonsten Anführungszeichen um die Keywords gestellt werden. Zusätzlich ist die Auswahl Write column header zu wählen (Abbildung 6, Ziffer 3), da Google Ads einen Titel für die darunter stehenden Keywords wünscht. Zuletzt muss im Reiter „Encoding“ noch UFT-8 gewählt werden, um mögliche Umlaute korrekt darzustellen (Abbildung 6, Ziffer 4).

Mit Ausführung des Knotens wird automatisch eine CSV-Datei erstellt und am gewählten Speicherort abgelegt.

## Attention, Attention! Es sind zu viele Keywords in der CSV-Datei

Der Keyword-Planer erlaubt nur eine Abfrage von maximal 10.000 Keywords pro CSV-Datei. Das klingt erst einmal sehr viel. Durch den Cross Join ergeben sich aber schnell sehr viele Kombinationen, die auch mal im sechsstelligen Bereich liegen können. Aber auch dafür hat KNIME natürlich eine Lösung parat!

## Das Chunking aka Keywords in Dateien aufteilen

Um sicherzustellen, dass die Abfrage in Google Ads problemlos läuft und ab einem Limit von 10.000 Einträgen in der CSV-Datei eine weitere Datei erstellt wird, werden die folgenden Knoten in den Workflow integriert (Abbildung 7). Hier wird auf das Prinzip des Chunkings zurückgegriffen, wobei die Daten in Chunks, kleine Teilbereiche, gruppiert werden.

### Schritt eins: CHUNK LOOP START

Zum Start kommt das Prinzip von Loopings ins Spiel. Hierbei wird eine Bedingung definiert, nach der der Loop immer wieder durchläuft, bis die Bedingung für alle Daten durchlaufen ist. Im vorliegenden Beispiel ist das notwendig, um die Keywords auf einzelne Dateien aufzuteilen. Die benötigte Bedingung: „Erstelle nach 10.000 Zeilen eine weitere CSV-Datei“. Bei 20.000 Zeilen wird der Loop zweimal durchlaufen und erstellt zwei CSV-Dateien.

Im CHUNK LOOP START wird deshalb das Limit von 10.000 Zeilen festgelegt. Somit wird das gesamte Datenset von oben nach unten durchgegangen und nach und nach werden die Daten in die einzelnen Chunks aufgeteilt.

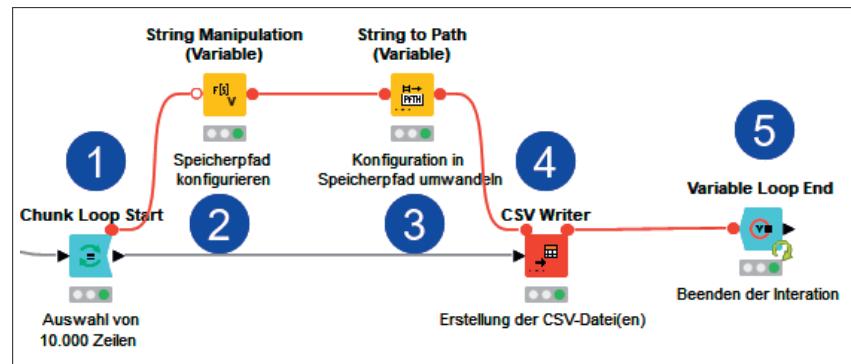


Abb. 7: Chunking der Zeilen in mehrere CSV-Dateien

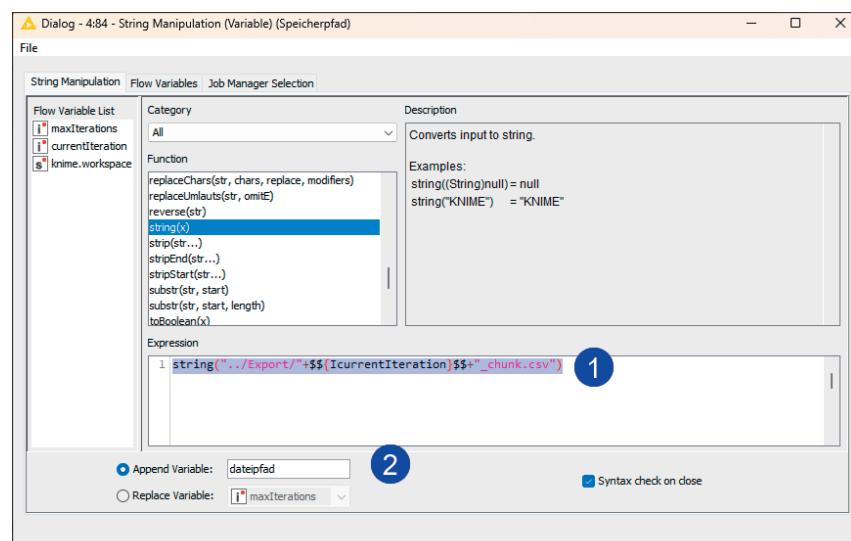


Abb. 8: Konfiguration des Dateipfads

### Schritt zwei: STRING MANIPULATION (VARIABLE)

Ähnlich wie der eben vorgestellte Knoten STRING MANIPULATION gibt es auch einen Knoten, der mit Variablen arbeitet. Dieser wird hier benötigt, um zu konfigurieren, wo die generierten CSV-Dateien abgespeichert und wie diese benannt werden sollen, da zum aktuellen Zeitpunkt nicht bekannt ist, wie viele Keywords vorliegen und wie viele CSV-Dateien mit je 10.000 Keywords erstellt werden müssen.

Hier wird deshalb eine Zeichenkette (= string) erstellt, die sich aus der Anzahl der Loopings und der Benennung des Pfads ergibt (Abbildung 8, Ziffer 1). Die Expression lautet:

**string(.../Export/\"+\${IcurrentIteration}\$\_chunk.csv")**

Wird die Expression so konfiguriert,

wird Folgendes erreicht:

» Erstellung eines Export-Ordners im

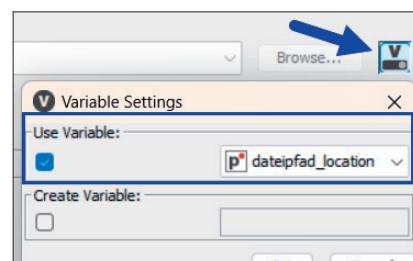


Abb. 9: Festlegung der Variable als Dateipfad

Ordner des Workflows

» Benennung der einzelnen CSV-Dateien als [Anzahl der Loops] + [\_chunk]

Zuletzt wird im Dialog noch Append Variable mit dem Namen „dateipfad“ gewählt (Abbildung 8, Ziffer 2).

### Schritt drei: STRING TO PATH

Dieser Knoten wird benötigt, um aus der neu generierten Zeichenkette nun einen interpretierbaren Dateipfad zu bekommen. Hierzu wird einfach die Zeichenkette mit dem festgelegten

Name	Änderungsdatum	Typ
0_chunk	28.08.2024 09:15	CSV-Datei
1_chunk	28.08.2024 09:15	CSV-Datei
2_chunk	28.08.2024 09:15	CSV-Datei
3_chunk	28.08.2024 09:15	CSV-Datei
4_chunk	28.08.2024 09:15	CSV-Datei
5_chunk	28.08.2024 09:15	CSV-Datei

Abb. 10: Ansicht der erstellten CSV-Dateien

Namen „dateipfad“ in das Feld Include geschoben. Ansonsten ist hier nichts einzustellen.

#### Schritt vier: CSV WRITER

Hier wird der Pfad für die Ablage und Benennung der Dateien anhand der vorher festgelegten Variablen definiert. Deshalb erscheint im Dialog die Warnmeldung „The path-parameter is controlled by a variable.“.

Zur Konfiguration wird das Symbol neben der Auswahl Browse geklickt. Es öffnet sich das Feld Variable Settings, in dem die im vorigen Knoten festgelegte Variable ausgewählt werden muss (Abbildung 9).

#### Schritt fünf: VARIABLE LOOP END

Damit der Loop ausgeführt werden kann, braucht es immer auch ein Ende des Loops. Im Knoten selbst muss nichts ausgewählt werden. Der Knoten wird einfach über Rechtsklick auf „EXECUTE“ ausgeführt. Während des Ausführens wird nun sichtbar, dass nun immer wieder alle Knoten des Loopings ausgeführt werden, bis alle Daten die Knoten durchlaufen haben. Nach 10.000 Zeilen werden die Knoten somit immer wieder durchlaufen.

Das Ergebnis sieht so aus: Es wurden CSV-Dateien mit der gewählten Benennung im gewählten Ordner abgelegt (Abbildung 10).

#### Abfragespaß in Google Ads

Der eigentliche Spaß beginnt nun, denn der Upload der CSV-Dateien in den Google-Ads-Keyword-Planer muss händisch erfolgen. Die Dateien werden nun einzeln abgefragt. Das Ergebnis kann nun einfach über die Export-Funktion zurück in einzelne CSV-Dateien erfolgen. Damit das Einlesen aller Dateien in KNIME gleich funktioniert, sollten alle Dateien in einem gemeinsamen Ordner abgelegt werden. Viel Spaß und bis gleich in KNIME!

#### Back to KNIME: Einlesen der CSV-Dateien

Zurück in KNIME braucht es nun einen CSV READER, um alle CSV-Dateien einzulesen. KNIME ermöglicht es hier über „Files in folder“ und „Browse“, alle Dateien eines bestimmten Ordners gleichzeitig einzulesen (Abbildung 11, Ziffer 1).

Weil der CSV-Export aus Google Ads, nett formuliert, eher suboptimal zur Weiterverbreitung ist, müssen hier noch folgende Anpassungen getroffen werden:

- » Column delimiter: \t (Abb. 11,2)
- » Reiter „Limit Rows“ → Skip first lines: 2 & Skip first data rows: 2

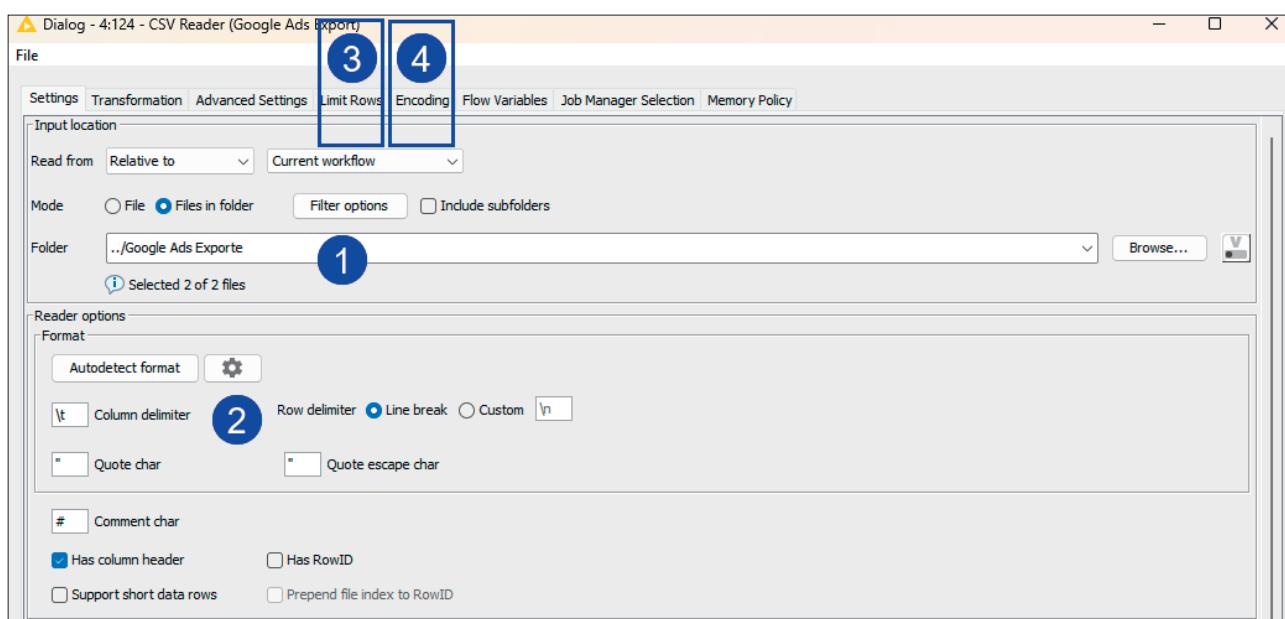


Abb. 11: Konfiguration des Knotens CSV READER für den Import

**HINWEIS**

Zunächst muss mithilfe der Dokumentation von SISTRIX die Abfrage-URL zusammengebaut werden, um das Suchvolumen je Keyword abzufragen. Für das Suchvolumen wird dazu die API-Funktion keyword.seo.traffic benötigt. Die API-URL sieht dazu so aus: [https://api.sistrix.com/keyword.seo.traffic?api\\_key=\[API\\_KEY\]&kw=\[keyword\]](https://api.sistrix.com/keyword.seo.traffic?api_key=[API_KEY]&kw=[keyword]). Benötigt werden daher der eigene API-Key und das gewünschte Keyword.

(Abbildung 11, Ziffer 3)

- » Reiter „Encoding“ → UTF-16 (Abbildung 11, Ziffer 4)

Nun werden alle CSV-Dateien untereinander eingelesen. Als Nächstes wird das Datenset gefiltert, sodass nun noch die relevanten Spalten „Keyword“ und „durchschnittliches Suchvolumen“ vorhanden sind. Dazu wird der COLUMN FILTER benötigt, in dem in die Includes nur diese beiden Spalten eingefügt werden.

### Die Aufbereitung zur Auswertung

Nun geht es schlussendlich noch um die eigentliche Auswertung. Um die Suchanfragen nach Produktnamen und nach transaktionalen Verb zu clustern, werden nun noch einmal die zugrunde liegenden Datensets benötigt. Es werden hier also die Daten aus dem Knoten der STRING MANIPULATION und die Daten aus dem hier zuletzt konfigurierten COLUMN FILTER kombiniert.

Dazu wird der JOINER benötigt, der ein Äquivalent zum SVERWEIS() in Excel ist. Hierbei wird als Referenzspalte die Spalte „Keyword“ benötigt, da diese in beiden Datensets vorliegt. Über die Auswahl wird deshalb im Top input „Keyword“ und auch im Bottom input „Keyword“ gewählt. Als Joining-Variante wird matching rows ausgewählt, da nur Spalten relevant sind, die in beiden Datensets vorliegen. Durch

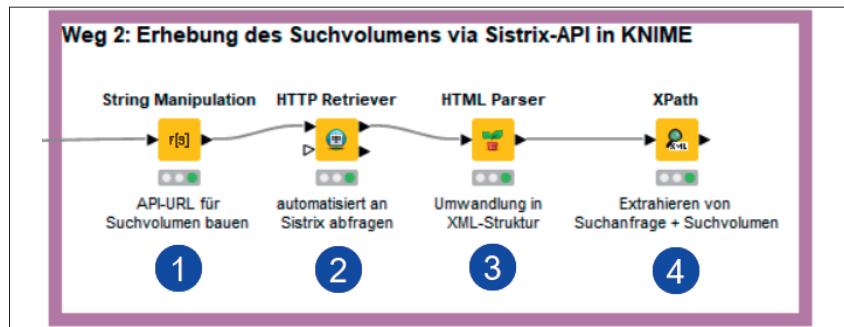


Abb. 12: Knoten für die Abfrage der SISTRIX-API

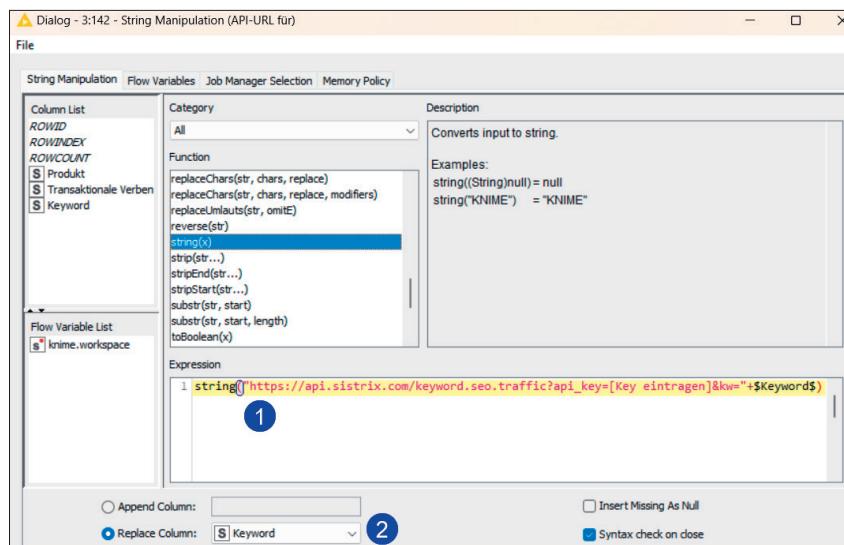


Abb. 13: Erstellung der API-URL zur Abfrage

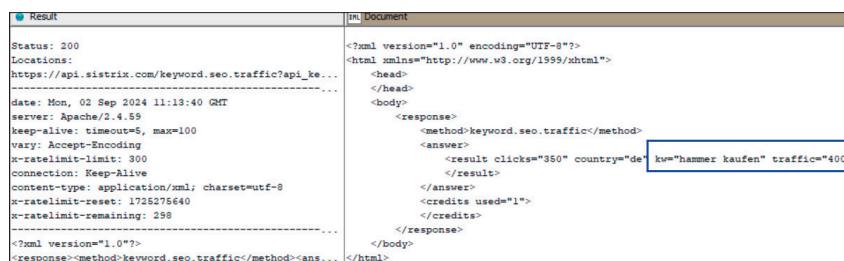


Abb. 14: Ergebnis im Knoten HTML PARSER

das Ausführen des Knotens werden die Datensets aneinandergeheftet.

Der nächste Schritt ist nun die Gruppierung der Daten.

### Weg zwei: Erhebung des Suchvolumens via SISTRIX-API in KNIME

Ein simplerer und auch noch schnellerer Weg, das Suchvolumen von Keywords zu erheben, ist eine Abfrage via API. Denn das funktioniert innerhalb der KNIME-Umgebung, ohne zwischendurch das Tool wechseln zu

müssen. Benötigt wird für diesen Weg jedoch der entsprechende Zugang zur SISTRIX-API.

**ACHTUNG:** Je nach SISTRIX-Account gibt es nur eine bestimmte Menge an Credits, die für API-Anfragen zur Verfügung stehen. Immer mal ein Auge darauf haben, nicht dass nach einer ungewollten Abfrage gleich alle Credits aufgebraucht sind. Pro Keyword-Abfrage wird ein Credit verbraucht. Für den Start empfiehlt es sich, den Ablauf erst einmal mit einem kleinen Set an Keywords auszuprobieren.

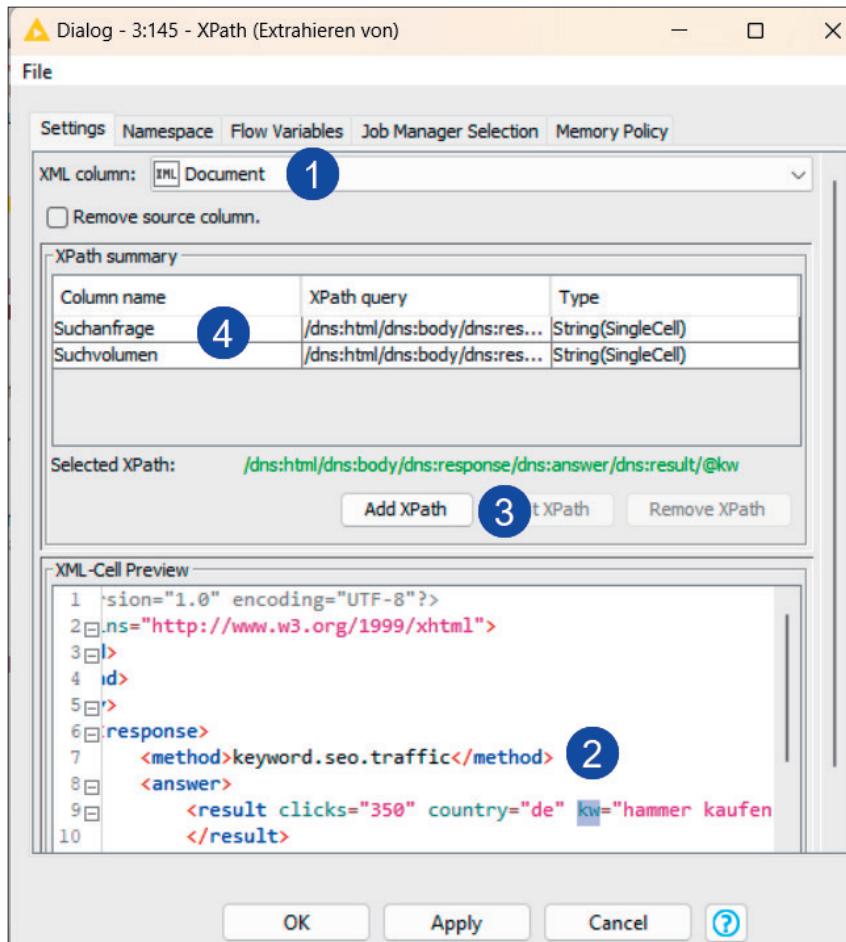


Abb. 15: Konfiguration im Knoten XPATH

Wie in Abbildung 3 gezeigt können die folgenden Knoten einfach an den ersten Knoten STRING MANIPULATION angehängt werden. Folgende Knoten sind, wie in Abbildung 12 zu sehen, für diesen Weg notwendig.

- 1. STRING MANIPULATION:** Um nicht für alle Keywords händisch eine eigene URL zu bauen, wird hier zunächst eine Zeichenkette gebaut, die mit der Spalte „Keyword“ aus dem vorangestellten Knoten des Datensets arbeitet. In der Expression (Abbildung 13, Ziffer 1) wird dazu Folgendes eingetragen:

```
string(„https://api.sistrix.com/keyword.seo.traffic?api_key=[Key eintragen]&kw=“+$Keyword$)
```

Unterhalb des Dialogs wird dann über Replace Column (Abbildung 13, Ziffer 2) die Keyword-Spalte durch die hier gebaute Zeichenkette ersetzt. Das Ergebnis ist, dass nun in jeder Zeile des Datensets eine Abfrage-URL steht, mit der das Suchvolumen je Keyword abgefragt werden kann.

- 2. HTTP RETRIEVER:** Um die Abfrage zu starten, wird im HTTP RETRIEVER über die Auswahl URL input angegeben, in welcher Spalte sich die Abfrage-URL befindet. Im vorliegenden Fall heißt die Spalte Keyword. Mit Ausführung des Knotens wird nun für jede URL der HTTP-Header abgefragt. Je nach Menge der Keywords kann das einen Moment dauern (Abbildung 12, Ziffer 2).

**3. HTML PARSER:** Mit dem nächsten Knoten wird die eben getätigten Abfrage nun in eine auslesbare XML-Datei umgewandelt. Dazu muss der Knoten einfach an den RETRIEVER angehängt werden. Im Dialogfeld wird im Drop-down-Menü als Input „Result“ gewählt, da die Inhalte dieser Spalte entsprechend geparsert werden sollen. Mit der Ausführung des Knotens wird an das Datenset eine weitere Spalte angehängt, die „Document“ heißt und die aus jeder Abfrage ein XML-Dokument zur Verfügung stellt (Abbildung 14).

**4. XPATH:** Um nun die relevanten Werte aus dem XML-Dokument zu extrahieren, muss auf die Abfragesprache XPath zurückgegriffen werden. Der gleichnamige Knoten XPATH unterstützt hierbei und muss zur Extraktion von Keyword und Suchvolumen wie folgt konfiguriert werden:

- » Zunächst wird über die Auswahl xml column ausgewählt, welche Spalte die XML-Dokumente enthält. Diese sollte im Default schon immer die Spalte „Document“ beinhalten (Abbildung 15, Ziffer 1).
- » In der Cell-Preview wird nun anhand der ersten Zeile ein Beispiel für das XML-Dokument angezeigt. Da alle nachfolgenden Dokumente gleich aufgebaut sind, kann anhand des Beispiels die Extraktion der benötigten Werte einfach erstellt werden. Mit Doppelklick auf die grünen Beschriftungen (hier: clicks, country, kw) kann jeweils ein Pfad ausgewählt werden, hinter dem sich der benötigte Wert befindet (Abbildung 15, Ziffer 2).
- » Wird mit Doppelklick je auf kw und auf traffic geklickt, erscheint im mittleren Bereich des Dialogs in grüner Schrift ein Pfad (Abbildung 15, Ziffer 3), der über „Add XPath“ hinzugefügt werden kann. Es öffnet sich nun ein

weiteres Feld, in dem nur ein Name vergeben werden muss. Danach wird eine neue Zeile in die Liste XPath summary aufgenommen (Abbildung 15, Ziffer 4).

- » Mit der Ausführung des Knotens ergeben sich nun für jede Zeile zwei neue Spalten, die das Keyword und das entsprechende Suchvolumen beinhalten.

### Bereinigung des Datensets

Nun werden mithilfe des Knotens COLUMN FILTER noch alle Spalten aus dem Datenset gefiltert, die für die Auswertung nicht benötigt werden. In den Includes werden deshalb nur noch die Spalten „Produktname“, „Verb“, „Keyword“ und „Suchvolumen“ beibehalten.

### Gruppierung und Auswertung der Daten

Die Gruppierung läuft nun wieder für beide Wege (CSV-Datei aus Google Ads und SISTRIX-Abfrage) gleich.

Um sich nun einen Überblick zu verschaffen, welche Produktnamen in Kombination mit dem Verb besonders relevant sind beziehungsweise welche Verben in Kombination mit dem Produktnamen das höchste Suchvolumen haben, werden die Daten nun gruppiert. Zum Gruppieren gibt es in KNIME den Knoten GROUP BY, der ähnlich wie eine Pivot-Konfiguration in Excel funktioniert.

Zur Gruppierung nach Verben wird im Dialog in den Group settings die Spalte „Transaktionale Verben“ eingefügt. Nun wird im Reiter „Manual aggregation“ noch ausgewählt, welche Werte an die einzelnen Gruppen geschrieben werden sollen. In diesem Beispiel soll das Suchvolumen je Gruppe aufsummiert dargestellt werden. Deshalb wird als Column „Suchvolumen“ und als Aggregation „Sum“ gewählt (Abbildung 16).

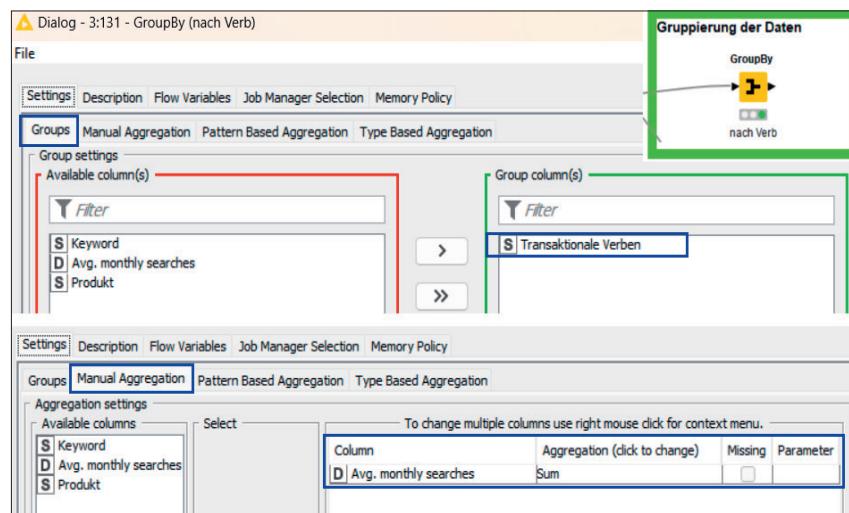


Abb. 16: Auswertung der Daten über den Knoten GROUP BY

Group table - 4:131 - GroupBy (nach Verb)			
File Edit Hilite Navigation View			
Table "default" - Rows: 7 Spec - Columns: 2 Properties Flow Variables			
Row ID	S Transaktionale Verben	D Avg. monthly...	
Row4	Kaufen	137,880	
Row1	Bestellen	12,220	
Row2	Einkaufen	100	
Row0	Auswählen	40	
Row3	Erwerben	?	
Row5	Reservieren	?	
Row6	Vorbestellen	?	

Abb. 17: Vorschau der Auswertung im Knoten GROUP BY

Das Ergebnis zeigt nun, dass die Kombinationen mit dem Verb „kaufen“ deutlich häufiger gesucht werden als mit dem Verb „bestellen“ oder „einkaufen“. Für Verben wie „reservieren“ gibt es keine Daten aus den Abfragen (Abbildung 17). Daraus lässt sich nun beispielsweise ableiten, dass eine Optimierung von Meta-Angaben einer Produktdetailseite dieses Produktsegments eher auf „kaufen“ als auf „einkaufen“ erfolgen sollte.

### Fazit

Der Workflow zeigt, wie sich Keyword-Kombinationen erstellen und abfragen lassen. Der Vorteil der Software ist hierbei wieder, dass jederzeit neue Dateien für die Kombination verwendet werden können, ohne dass der Workflow angepasst werden muss. Zusätzlich ist der große Vorteil, dass

die Keyword-Gruppen erhalten bleiben, wodurch schlussendlich eine Gruppierung der Keywords und somit eine hilfreiche Auswertung der Recherche erfolgen kann.

Sogenannte Keyword-Mixer-Tools gibt es im Internet kostenfrei zuhauf, jedoch ergibt sich hier keine Möglichkeit der Gruppierung und Auswertung.

Die Workflows wurden zur besseren Übersicht nach den beiden beschriebenen Wegen aufgeteilt und sind unter [einfach.st/knime86](http://einfach.st/knime86) abrufbar und per Drag-and-drop einfach in der eigenen KNIME-Umgebung zu verwenden. In diesem Sinne viel Spaß mit der Arbeit in KNIME! ¶