



# MIT SEARCH CONSOLE & KNIME: CONTENT-DAUERBRENNER UND LAHME ENTEN SICHTBAR MACHEN

Michael Göpfert

Die Analyse von großen Dokumenten und Keywordlisten ist oft mühselig und echter Erkenntnisgewinn schwierig. Eine Analyse von N-Grammen kann dabei eine völlig neue Perspektive eröffnen, Zusammenhänge aufzeigen und Potenziale veranschaulichen – sowohl bei der Wettbewerbsanalyse als auch bei der eigenen Website und der Keyword-Recherche. Michael Göpfert zeigt Ihnen, wie Sie mit dem kostenlosen und wirklich von jedermann bedienbaren Tool KNIME recht einfach Antworten auf solche Problemstellungen erhalten.

Ein häufig gehörtes Versprechen von SEOs lautet: Ist ein hochwertiges Stück Content erst einmal erstellt, bringt dieses über einen langen Zeitraum kontinuierlich Traffic über Suchmaschinen. Doch welche Inhalte einer Website können dieses Versprechen wirklich einlösen?

Analysiert wird diese Fragestellung häufig wie folgt:

- » Betrachtung der URLs mit den meisten Klicks: Einzelereignisse, die viel Traffic bringen, oder Saisonalitäten werden dabei nicht berücksichtigt.
- » Betrachtung einzelner URLs im Zeitverlauf: Dabei werden zwar saisonale Effekte sichtbar, die Analyse konzentriert sich jedoch nur auf einzelne URLs.

Beide Herangehensweisen haben ihre Stärken (die Seiten mit dem meisten Traffic werden sichtbar), haben aber auch blinde Flecke. Häufig finden sich auf einer Website Seiten, die im Schatten der Top-Performer kontinuierlich Traffic bringen. Um das Identifizieren solcher Evergreen-URLs soll es in diesem Beitrag gehen.

Dabei wird ermittelt, an wie vielen Tagen im Jahr jede URL einer Website Klicks über die Google-Webuche bekommt. Die Analyse beantwortet unter anderem folgende Fragestellungen:

- » Welche Seiten sind „echter“ Evergreen-Content und bringen der Website an vielen Tagen im Jahr Traffic?
- » Welche der Seiten bekommen schon län-

## DER AUTOR



**Michael Göpfert** arbeitet gerne mit Rohdaten, um diese in maßgeschneiderten Analysen für seine Kunden aufzubereiten.

Foto: mikedam / gettyimages.de

ger keinen Klick mehr und sollten deshalb aktualisiert oder gelöscht werden?

Die Daten dazu liefert die Google Search Console (GSC), die Auswertung der Daten erfolgt über das kostenlose Datenanalyse-Tool KNIME.

Da bei dieser Art der Analyse sehr viele Daten anfallen, reicht es nicht, die Daten aus dem Leistungsbericht der GSC zu exportieren, da die Exporte auf 1.000 Zeilen limitiert sind. Um diese Beschränkungen zu umgehen, können die Daten über die API der Search Console geholt werden.

Leider bietet KNIME keine native Integration der API der GSC. Durch einen kleinen Umweg über Google Sheets können die Daten aber trotzdem via API gewonnen werden. Das Google-Sheets-Add-on „Search Analytics for Sheets“ ermöglicht es GSC-Usern, die Daten aus dem Leistungsbericht direkt in eine Google-Tabelle zu importieren.

### Export der Daten aus der Google Search Console

Nach der Installation von Search Analytics for Sheets lässt sich die Erweiterung direkt in Google Sheets unter dem Menüpunkt „Erweiterungen“ aufrufen. Anschließend müssen noch folgende Einstellungen (Abb. 2) vorgenommen werden:

- » Verified Seite: Die GSC Property auswählen, die abgefragt werden soll
- » Date Range: einen Zeitraum von 365 Tagen auswählen
- » Group By: Date und Page
- » Haken setzen bei Include fresh data

Alle anderen Einstellungen können unverändert bleiben. Anschließend kann die Tabelle als CSV exportiert werden.

### Import der Daten in KNIME

Liegen die Daten als CSV-Datei auf der Festplatte, können diese in KNIME

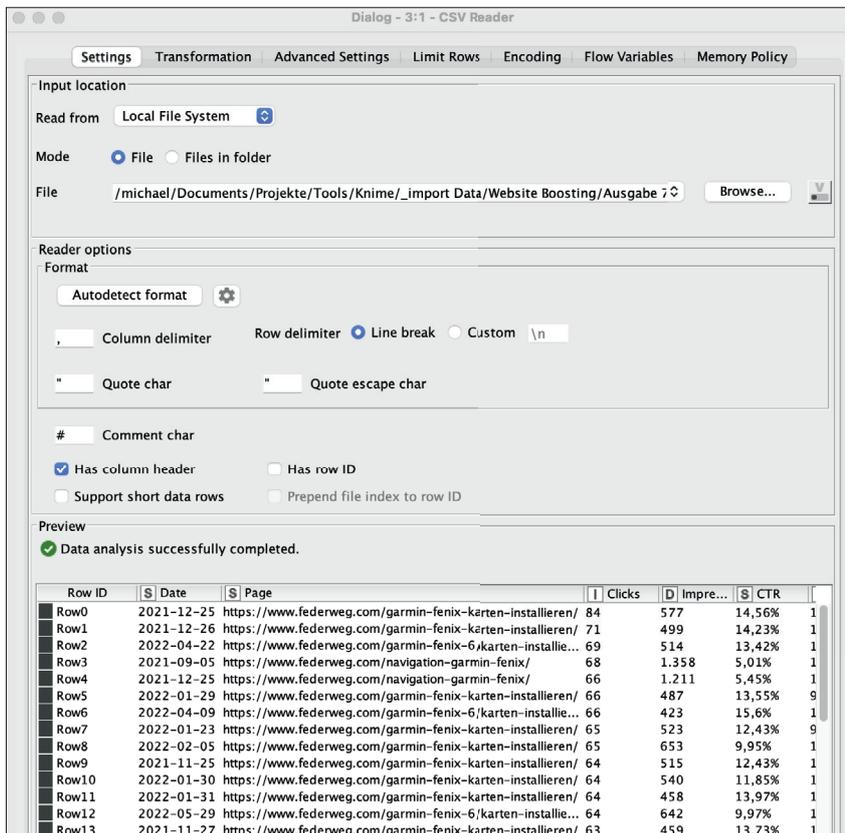


Abb.1: Import der CSV-Datei mit dem CSV-Reader in KNIME

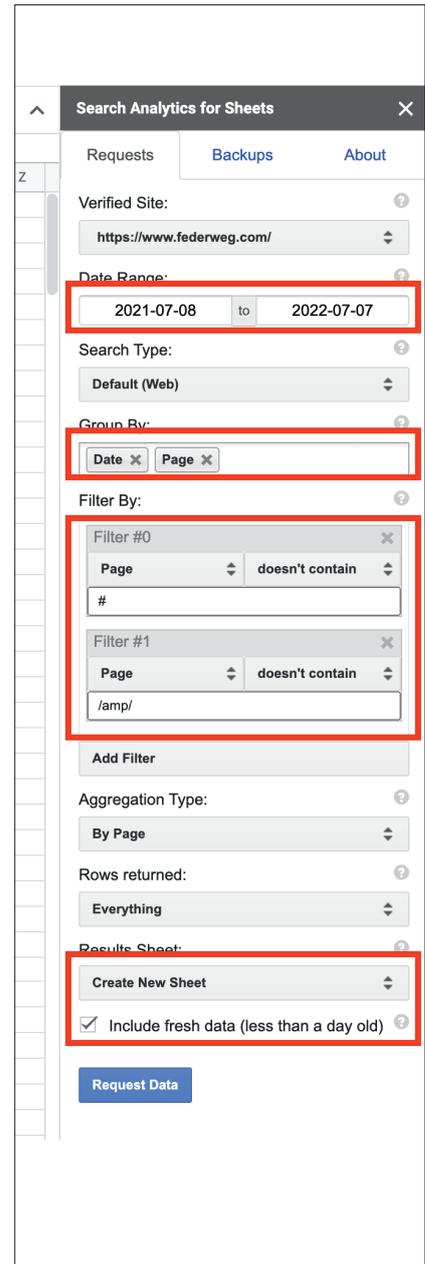


Abb. 2: Die wichtigsten Einstellungen in Search Analytics for Sheets

### HINWEIS

Bei dieser Art der Abfrage fallen große Mengen an Daten an, die das Zeilenlimit von Google Sheets überschreiten können. Bei größeren Websites kann daher in Search Analytics for Sheets ein Filter gesetzt werden, damit nur ein Teilbereich (z. B. ein bestimmtes Verzeichnis) der Website abgefragt wird. Auch können kanonisierte Seiten (Seiten mit # in der URL oder AMP-Seiten) direkt bei der Abfrage über „Filter By“ herausgefiltert werden.

**TIPP**

Wie Sie an das kostenlose Tool KNIME kommen und wie es prinzipiell funktioniert, finden Sie in der Ausgabe 53 oder online frei als HTML oder PDF unter <http://einfach.st/knime53>.



importiert werden. Zunächst sollte ein neuer Workflow in KNIME erstellt werden. Anschließend kann aus dem Node Repository die Node CSV Reader auf die Arbeitsfläche gezogen werden.

Mit einem Doppelklick auf die Node öffnet sich ihr Konfigurationsmenü. Dort kann die CSV-Datei mit den GSC-Daten ausgewählt werden. Der CSV-Reader erkennt in der Regel von selbst, wie die CSV-Datei aufgebaut ist, sodass keine weiteren Einstellungen vorgenommen werden müssen.

Durch einen Rechtsklick auf die Node und das Auswählen des Menüpunkts „Execute“ lässt sich die Node ausführen, und die Daten werden importiert. Durch erneutes Aufrufen des Kontextmenüs der Node per Rechtsklick und Auswahl des Punkts „File Table“ lässt sich die importierte Tabelle anzeigen. Sie sollte die Spalten „Date“, „URL“, „Clicks“, „Impressions“, „CTR“ und „Position“ enthalten (Abb. 3).

**Datumsfunktionen in KNIME**

Um mit Daten in Form von Tagen, Monaten oder Jahren zu arbeiten, muss zunächst der Datentyp der Spalte „Date“ von String (= Text) in ein Datum umgewandelt werden. Diese Prozedur ist vielen aus Excel bereits bekannt.

In KNIME gibt es dazu die Node String to Date&Time. Diese kann eben-

S	Date	S	Page	I	Clicks	D	Impre...	S	CTR	S	Position
2021-12-25	https://www.federweg.com/garmin-fenix-karten-installieren/	84	577	14,56%	10,2						
2021-12-26	https://www.federweg.com/garmin-fenix-karten-installieren/	71	499	14,23%	12,2						
2022-04-22	https://www.federweg.com/garmin-fenix-6/karten-installe...	69	514	13,42%	11,9						
2021-09-05	https://www.federweg.com/navigation-garmin-fenix/	68	1.358	5,01%	12,2						
2021-12-25	https://www.federweg.com/navigation-garmin-fenix/	66	1.211	5,45%	11,3						
2022-01-29	https://www.federweg.com/garmin-fenix-karten-installieren/	66	487	13,55%	9,9						
2022-04-09	https://www.federweg.com/garmin-fenix-6/karten-installe...	66	423	15,6%	13,2						
2022-01-23	https://www.federweg.com/garmin-fenix-karten-installieren/	65	523	12,43%	9,8						
2022-02-05	https://www.federweg.com/garmin-fenix-karten-installieren/	65	653	9,95%	19,3						
2021-11-25	https://www.federweg.com/garmin-fenix-karten-installieren/	64	515	12,43%	12,5						
2022-01-30	https://www.federweg.com/garmin-fenix-karten-installieren/	64	540	11,85%	10,7						
2022-01-31	https://www.federweg.com/garmin-fenix-karten-installieren/	64	458	13,97%	11,7						
2022-05-29	https://www.federweg.com/garmin-fenix-6/karten-installe...	64	642	9,97%	16,9						
2021-11-27	https://www.federweg.com/garmin-fenix-karten-installieren/	63	459	13,73%	13,6						
2022-03-27	https://www.federweg.com/garmin-fenix-6/karten-installe...	63	535	11,78%	16,2						
2022-01-23	https://www.federweg.com/navigation-garmin-fenix/	61	1.174	5,2%	12,8						
2022-05-22	https://www.federweg.com/garmin-fenix-6/karten-installe...	61	533	11,44%	15,0						
2021-11-26	https://www.federweg.com/garmin-fenix-karten-installieren/	60	500	12%	13,3						
2022-01-21	https://www.federweg.com/garmin-fenix-karten-installieren/	60	486	12,35%	13,3						

Abb.3: Die importierte CSV-Datei

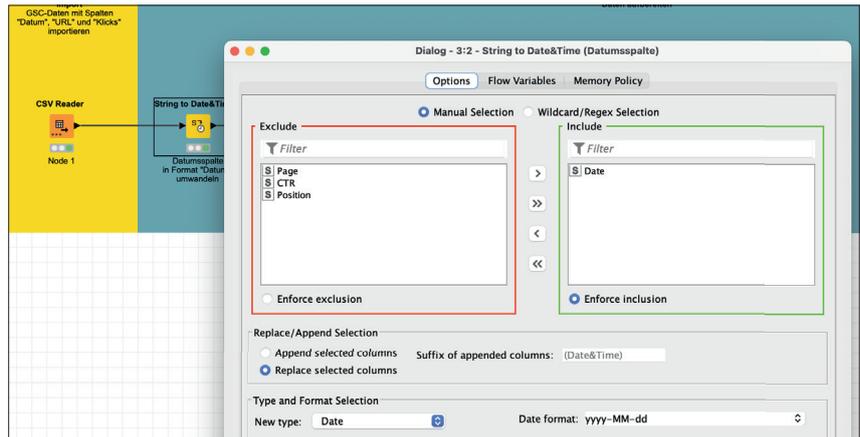


Abb.4: Änderung des Datentyps in Date&Time

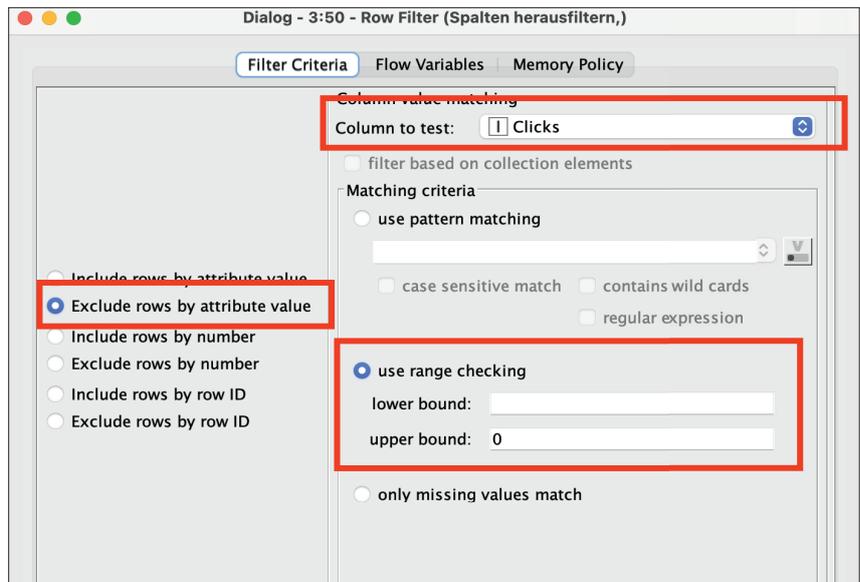


Abb. 5: Herausfiltern der Zeilen, die 0 Klicks beinhalten

falls vom Repository auf die Arbeitsfläche gezogen und an den schwarzen Dreiecken mit dem CSV-Reader verbunden werden.

In der Konfiguration dieser Node muss zunächst nur die Spalte ausgewählt werden, die das Datum enthält.

Alle anderen Spalten sollten in den Bereich „Exclude“ verschoben werden (Abb. 4).

Mithilfe des Buttons Guess data type and format kann die Node erkennen, in welchem Format das Datum in der Spalte „Date“ aufgebaut ist, und die

### Herausgeber & Chefredakteur (verantwortlich):

Mario Fischer  
E-Mail: redaktion@websiteboosting.com

### Autoren dieser Ausgabe:

Dr. Martin Bahr, Alexander Beck,  
Britta Behrens, Torsten Beyer,  
Michael Göpfert, Joanna Hengstebeck,  
Thomas Kaiser, Max Lamprecht,  
Rico Melzer, Sarah Weitnauer

### Anzeigenleitung:

Markus Lutz  
E-Mail: anzeigenleitung@websiteboosting.com

### Art Direction, Layout/Produktion:

Kai Neugebauer

### Lektorat:

Bärbel Philipp, textperlen.de,  
Ursula Wenke, www.ektorat-wenke.de

### Fotos & Illustrationen:

Website Boosting / GettyImages

### Druck:

Vogel Druck und Medienservice GmbH  
Leibnizstr. 5, 97204 Höchberg

### Vertrieb:

PressUp GmbH  
Postfach 70 13 11  
22013 Hamburg  
E-Mail: websiteboosting@pressup.de

### Abonnement:

Website Boosting Aboservice  
PressUp GmbH  
Postfach 70 13 11  
22013 Hamburg  
Tel. 040 / 38 6666 - 342  
Fax: 040 / 38 6666 - 299  
E-Mail: websiteboosting@pressup.de

Erscheinungsweise: 6 x jährlich  
Bezugspreis: Einzelheft: 11,80€

Bezugspreis Inland jährlich 62,00€ inkl. Versand  
Bezugspreis Ausland jährlich 70,80€  
inkl. Versand

Studenten im Inland erhalten gegen Vorlage einer  
Immatrikulationsbescheinigung einen  
Preisvorteil – Details finden Sie auf der Website.

### Verlagsleitung:

Michael Müßig  
Tel: +49 931 / 26 038 04,  
verlag@websiteboosting.com

### Anschrift des Verlages

Hotspot Verlag GmbH  
Obere Landwehr 4a, 97204 Höchberg  
Tel: + 49 931 / 26 038 04  
Fax: +49 931 / 26 038 05  
E-Mail: verlag@hotspotverlag.de  
www.hotspotverlag.de

### Geschäftsführung:

Kai Neugebauer

Die Inhaber- und Beteiligungsverhältnisse  
lauten wie folgt: Gesellschafter zu 100%  
ist die Webvalue Holding GmbH

ISSN: 2191-6241

Für unverlangt eingereichte Texte und Daten kann keine Haftung übernommen werden. Sämtliche Veröffentlichungen in Website Boosting erfolgen ohne Berücksichtigung eines eventuellen Patentschutzes. Markennamen werden ohne Gewährleistung einer freien Verwendung benutzt. Trotz sorgfältiger Recherche kann für die Richtigkeit des Inhalts keine Haftung übernommen werden. Namentlich gekennzeichnete Artikel geben nicht unbedingt die Meinung der Redaktion wider.

**Jambo Bukoba e.V.** ist eine gemeinnützige Organisation mit Sitz in München. Wir unterstützen und fördern Projekte unserer Partnerorganisation Jambo for Development im Nordwesten Tansanias in den Bereichen **Bildung, Gesundheit und Gleichberechtigung**. Sowohl wir als Verein als auch die Hilfsprojekte in Tansania folgen den Prinzipien der Hilfe zur Selbsthilfe und Sport für Entwicklung für eine nachhaltige und ganzheitliche **Förderung von Kindern und Jugendlichen in Afrika**.



Du willst uns und unsere Arbeit unterstützen? Dann schau Dir auf [www.jambobukoba.com](http://www.jambobukoba.com) unser einzigartiges Konzept und unsere individuellen Projekte an und finde heraus, wie Deine Spende Kindern in Tansania die Chance auf eine bessere Zukunft ermöglicht.

## Mach mit uns gemeinsam Kinder in Tansania stark!



proudly supported by

**121WATT**  
we share knowledge

SEO | SEM | E-COMMERCE | USABILITY | SZENE | TIPPS & TOOLS  
**WEBSITE BOOSTING**

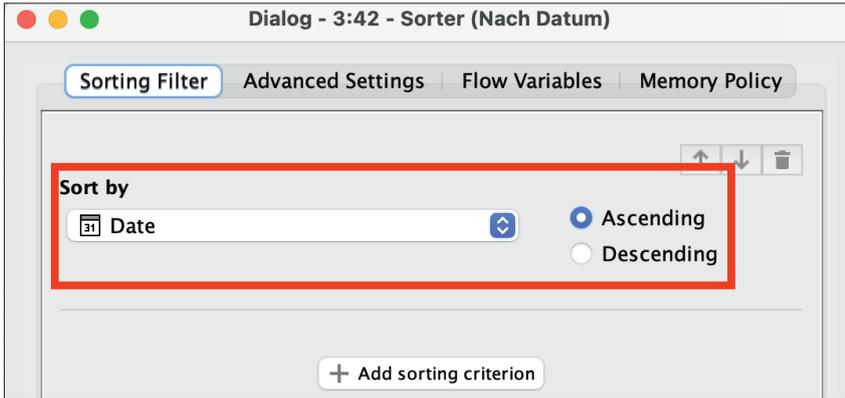


Abb.6: Nach Datum absteigend sortieren

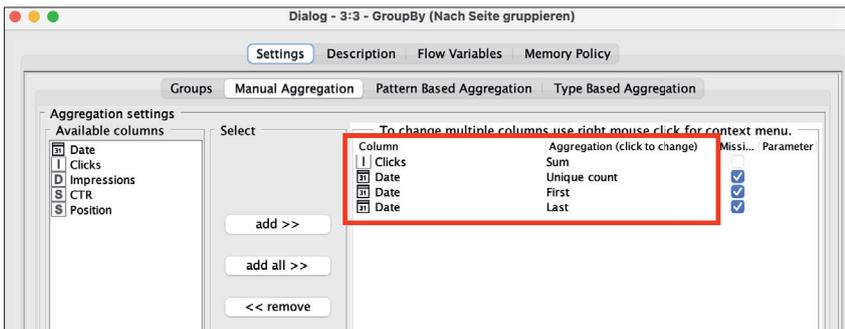


Abb. 7: Gruppieren und neue Spalten erzeugen

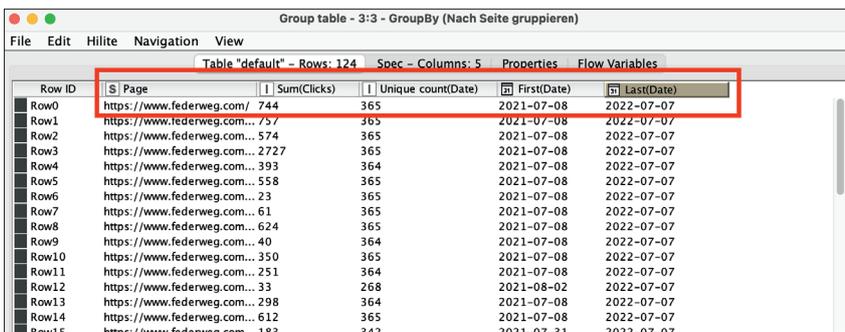


Abb. 8: Die gruppierte Tabelle

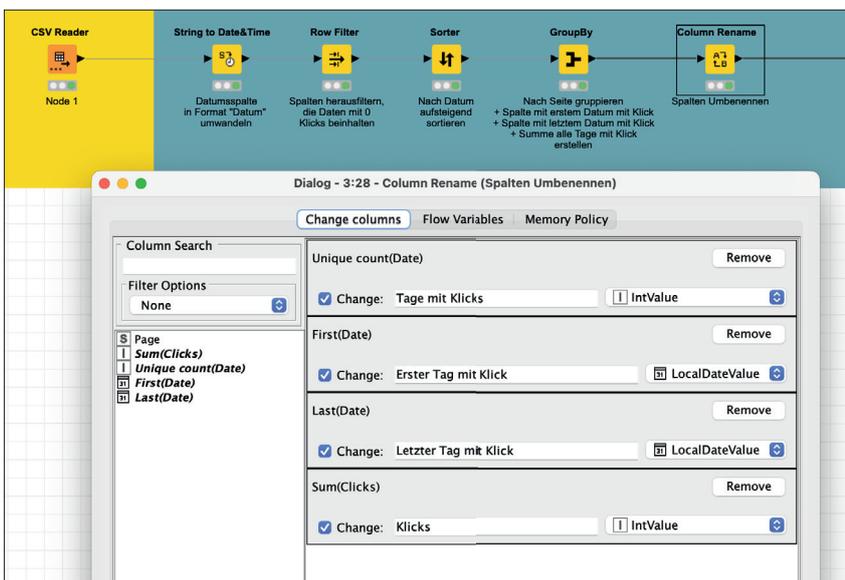


Abb. 9: Umbenennung der Spalten

Spalte in den Datentyp „Date&Time“ umwandeln.

Im nächsten Schritt werden mit der Node Row Filter alle Zeilen herausgefiltert, die in der Spalte „Clicks“ eine 0 stehen haben (Abb. 5). Schließlich bezieht sich die Analyse ausschließlich auf URLs, die auch Klicks bekommen.

Nun wird die Tabelle so sortiert, dass das älteste Datum am Anfang der Tabelle steht. Dazu gibt es die Node Sorter. Hier werden in der Konfiguration die Spalte „Date“ und die Sortiermethode „Ascending“ gewählt (Abb. 6).

### Gruppieren der Daten für erste Insights

Aktuell enthält die Tabelle für jedes Datum und jede URL eine Zeile. Mithilfe der Node GroupBy wird die Tabelle nun auf eine Zeile pro URL reduziert. Die Werte aus der Spalte „Date“ werden genutzt, um zu ermitteln, an wie vielen Tagen eine jede URL Klicks bekommen hat.

Dazu wird im Tab „Groups“ in der Konfiguration der Node GroupBy die Spalte „Page“ ausgewählt. Anschließend wird im Tab „Manual Aggregation“ die Spalte „Clicks“ hinzugefügt und im Feld „Aggregation“ wird Sum ausgewählt. Das bewirkt, dass zu jeder URL die Summe ihrer Klicks ermittelt wird.

Zusätzlich wird dreimal die Spalte „Date“ hinzugefügt und im Feld „Aggregation“ wird jeweils Folgendes (Abb. 7) ausgewählt:

- » **Unique Count:** Dies ermittelt, an wie vielen Tagen eine URL Klicks bekommen hat. Jeder unique Wert in der Spalte wird ermittelt und gezählt.
- » **First:** Diese Einstellung erzeugt eine neue Spalte mit dem Datum, an dem die URL zum ersten Mal einen Klick bekommen hat.
- » **Last:** Diese Einstellung erzeugt eine neue Spalte mit dem Datum, an dem die URL zum letzten Mal einen Klick bekommen hat.

Für jede der getroffenen Einstellungen erzeugt die Node eine neue Spalte mit den Werten, die sich aus den getroffenen Einstellungen ergeben. Die Tabelle besteht nun aus den folgenden Spalten (Abb. 8):

- » Page: die URL
- » Sum(Clicks): die Summe aller Klicks der jeweiligen URL
- » Unique Count(Date): Anzahl an Tagen, an denen die URL mindestens einen Klick bekommen hat
- » First(Date): das Datum, an dem die URL ihren ersten Klick bekommen hat. Die Node nimmt hier von oben nach unten den ersten Wert, den sie für jede URL in der Spalte „Date“ findet.
- » Last(Date): das Datum, an dem die URL das letzte Mal einen Klick bekommen hat. Die Node nimmt hier von oben nach unten den letzten Wert, den sie für jede URL in der Spalte „Date“ findet. Deshalb ist die absteigende Sortierung aus dem vorherigen Schritt ganz entscheidend.

Nun sind schon erste Ergebnisse sichtbar: In der Spalte Unique Count(Date) zeigt sich, an wie vielen Tagen jede URL mindestens einen Klick bekommen hat.

Um die Namen der Spalten etwas selbsterklärender zu machen, können diese mit der Node Column Rename umbenannt werden.

### Berechnen eines „Evergreen Scores“

Beim Betrachten der Tabelle zeigt sich jedoch ein Problem: Ist eine URL seit weniger als einem Jahr live, kann diese noch so häufig Traffic bekommen, sie wird vermutlich nicht so viele Tage mit Klicks haben wie eine ältere URL. Deshalb ist es sinnvoll, einen Wert zu ermitteln, der auch diesen URLs gerecht wird. Dies kann erreicht werden, indem berechnet wird, an wie viel Prozent der Tage Klicks auf eine

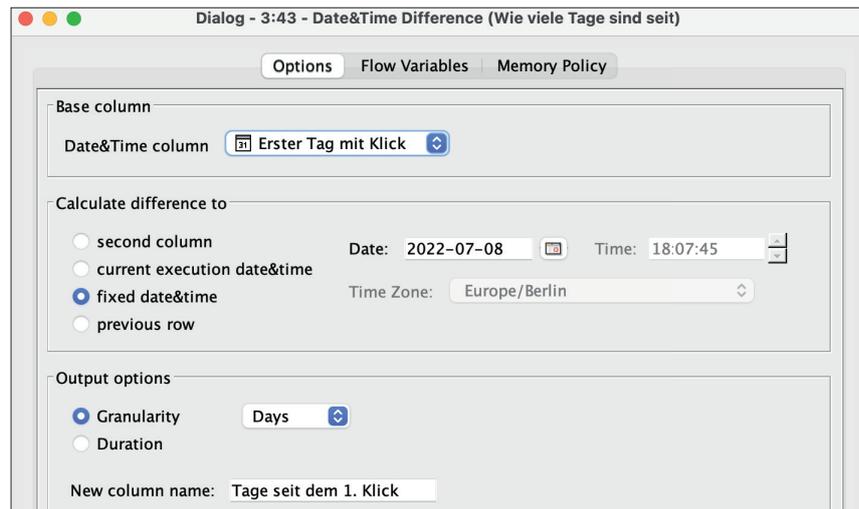


Abb. 10: Ermitteln, wie viele Tage seit dem ersten Klick auf eine URL vergangen sind

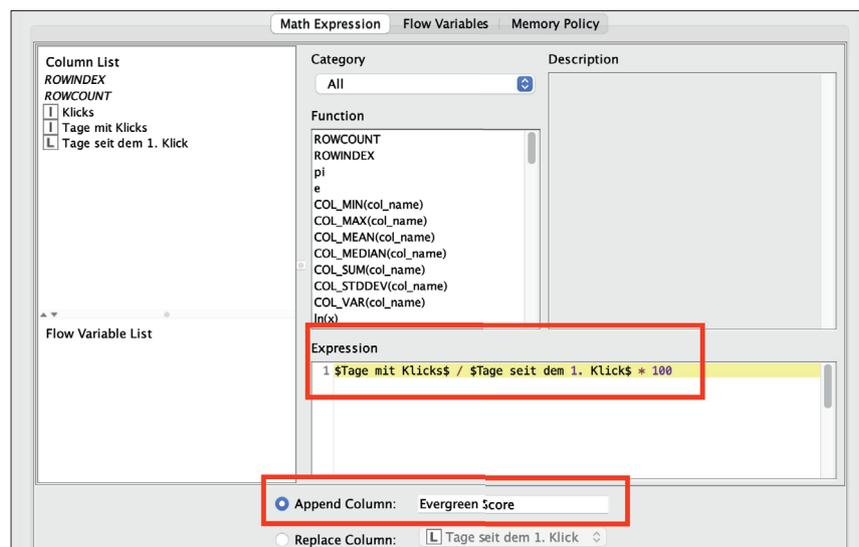


Abb. 11: Berechnen des „Evergreen Scores“

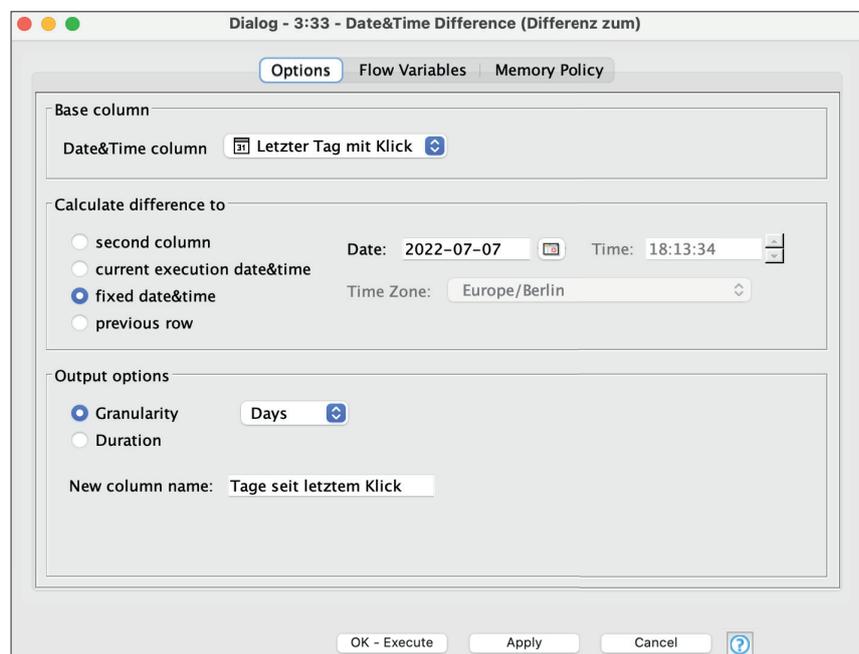


Abb. 12: Ermitteln, wie viele Tage seit dem letzten Klick vergangen sind

Show 10 entries

Search:

RowID	Page	Klicks	Tage mit Klicks	Erster Tag mit Klick	Letzter Tag mit Klick	Tage seit dem 1. Klick	Tage seit letztem Klick	Evergreen Score
Row0	https://www.federweg.com/	744	315	2021-07-08	2022-07-07	365	0	86.3013698630137
Row1	https://www.federweg.com/3-laender-enduro-trails/	757	243	2021-07-08	2022-07-07	365	0	66.57534246575342
Row2	https://www.federweg.com/albrauf-mtb/	574	271	2021-07-08	2022-07-07	365	0	74.24657534246575
Row3	https://www.federweg.com/alpencross-packliste/	2727	349	2021-07-08	2022-07-07	365	0	95.61643835616438
Row4	https://www.federweg.com/altlacher-hochkopf-trail/	393	180	2021-07-08	2022-07-06	365	1	49.31506849315068
Row5	https://www.federweg.com/andechser-holy-trails/	558	250	2021-07-08	2022-07-07	365	0	68.4931506849315
Row6	https://www.federweg.com/ausruistung/	23	23	2021-07-16	2022-06-24	357	13	6.442577030812324
Row7	https://www.federweg.com/ausruistung/mountainbike-bekleidung/	61	55	2021-07-11	2022-06-19	362	18	15.193370165745856
Row8	https://www.federweg.com/ausruistung/mountainbike-schuhe/	624	283	2021-07-08	2022-07-07	365	0	77.53424657534246
Row9	https://www.federweg.com/axa-greenline-30/	40	35	2021-07-18	2022-06-06	355	31	9.859154929577464

Showing 1 to 10 of 119 entries

Previous 1 2 3 4 5 ... 12 Next

Abb. 13: Die finale Tabelle

URL kommen. Dieser Wert ist unabhängig davon, wie alt eine URL ist.

Dazu muss zunächst ermittelt werden, wie lange der Tag mit dem ersten Klick zurückliegt. Dazu bietet sich die Node Date&Time Difference an. Um den Wert zu ermitteln, muss die Node wie in Abbildung 10 konfiguriert werden.

Das eingestellte Datum im Bereich „Calculate difference to“ sollte dabei dem letzten Tag des Zeitraums, der für den Export der Daten aus der GSC ausgewählt wurde, plus einen Tag entsprechen. Warum plus einen Tag? Ganz einfach: Die Datumsauswahl in der Node „Date&Time Difference“ berücksichtigt alle Tage bis zum gewählten Datum, aber nicht einschließlich.

Nun kann mit der Node „Math Formula“ die Spalte „Tage mit Klicks“ durch die Spalte „Tage seit dem ersten Klick“ geteilt und mit 100 multipliziert werden (Abb. 11). Diese Werte werden mit der Einstellung „Append Column“ in eine neue Spalte mit dem Namen „Evergreen Score“ geschrieben.

Der sogenannte „Evergreen Score“ zeigt nun an, an wie viel Prozent der Tage, seitdem die URL das erste Mal einen Klick bekommen hat, die Seite weitere Klicks bekommen hat.



Abb. 14: Ziehen Sie einfach den fertigen Workflow direkt in KNIME

Ebenfalls hilfreich ist die Angabe, wie viele Tage vergangen sind, seitdem eine URL das letzte Mal einen Klick bekommen hat. Damit lassen sich Seiten aufdecken, die schon länger keinen Traffic mehr bekommen und eventuell aktualisiert oder gelöscht werden sollten.

Das Vorgehen ist hier ähnlich wie bei der Ermittlung des Zeitraums seit dem ersten Klick, nur dass in der Konfiguration die Spalte „Letzter Tag mit Klick“ ausgewählt werden muss (Abb. 12). Auf das Erweitern der Datumsangabe um einen Tag kann hier verzichtet werden.

Um die Spalten noch in eine attraktivere Reihenfolge zu bringen, können diese mit der Node Column Resorter nach Belieben umsortiert werden.

Zum Schluss kann die Tabelle noch mit der Node Table View visualisiert werden. So entsteht eine filter- und

sortierbare Tabelle mit den erhobenen Daten, die mit der Node Excel-Writer auch in das bekannte Excel-Format exportiert werden kann.

**Fazit**

Die Kombination aus Search-Console-Daten und KNIME ermöglicht spannende neue Insights, ganz ohne programmieren zu müssen. Oder wussten Sie, wie viele Ihrer URLs seit über 100 Tagen keinen Traffic mehr bekommen?

Probieren Sie es also am besten selbst mal aus. Wie immer kann der fertige Workflow auch heruntergeladen werden. Dazu müssen Sie einfach nur mit der Maus das gelbe viereckige Symbol rechts oben in der Kopfzeile der Seite auf die Arbeitsfläche von KNIME ziehen (Abb. 14): <http://einfach.st/knime9>. ¶