



Google BigQuery und Tableau: Best Practices

Einleitung

Mithilfe von Tableau und Google BigQuery lassen sich über eine benutzerfreundliche grafische Oberfläche innerhalb kürzester Zeit riesige Datenmengen analysieren und komplexe Fragen beantworten. Zusammen bieten Ihnen diese beiden Tools folgende Möglichkeiten:

- Normalen Business-Anwendern mit der Leistungsfähigkeit von Google BigQuery schnelle und ... interaktive Analysen an die Hand geben
- Innerhalb weniger Sekunden Milliarden von Zeilen visuell analysieren, ohne serverseitige Verwaltung und ohne Programmieraufwand
- In Minutenschnelle beeindruckende Dashboards erstellen, die mit Ihren Google BigQuery-Daten verbunden sind und Ihr Unternehmen immer auf dem Laufenden halten
- Berichte und Erkenntnisse mit Tableau Server und Tableau Online im Internet freigeben, sodass jeder von einem beliebigen Gerät aus darauf zugreifen kann
- Die Cloud-Flexibilität von Google BigQuery mit der beeindruckenden Geschwindigkeit von Tableau kombinieren und den Projektwert schneller erkennen

Durch die Optimierung beider Technologien lassen sich deutliche Leistungsverbesserungen erzielen, Designzyklen verkürzen sowie Benutzern und Unternehmen zu mehr Erfolg verhelfen. In diesem Whitepaper erläutern wir, wie Sie Datenmodellierung und Abfrageformulierung optimieren, um eine maximale Reaktionsfähigkeit der Visualisierungen zu gewährleisten. Wir werden uns auch Verfahren widmen, mit denen sich bei der gemeinsamen Nutzung von Tableau und BigQuery die beste Kosteneffizienz erzielen lässt.

Verfasser

Pierce Young, Product Manager, Tableau

Vaidy Krishnan, Senior Product Manager, Tableau

Riley Maris, Senior Product Marketing Specialist, Tableau

Babu Prasad Elumala, Solutions Engineer, Google

Seth Hollyman Technical Program Manager, Google

Tino Tereshko, Enterprise Solutions Engineer, Google

Mike Graboski, Solutions Engineer, Google

Inhalt

Technologieübersicht	4
Google BigQuery	4
Tableau	5
Best Practices für Performance: Tableau	6
Leistungsaufzeichnung	6
Kontextfilter.....	7
Sets und Gruppen.....	8
Filter zuerst hinzufügen	8
Deaktivieren der automatischen Aktualisierungen.....	9
Auf Warnungen achten	9
Optimieren paralleler Abfragen.....	10
Best Practices für Kosten und Leistung: Google BigQuery	10
Denormalisierung und Vorabverknüpfung	10
Horizontale Fragmentierung von Tabellen nach Datum	11
Angabe einer Zieltabelle beim Ausführen vieler ähnlicher Abfragen	12
Google BigQuery ML-Modellresultate mit Tableau visualisieren	12
Fallstudie: Top-Tipps für Selfservice-Analytics mit Tableau und Google BigQuery	
von zulily	13
Fazit	13
Über Tableau	14
Weitere Materialien	14

Technologieübersicht

Google BigQuery

BigQuery kann innerhalb von Sekunden mehrere Petabyte an Daten mit SQL-Befehlen verarbeiten, ohne dass hierfür eine Feinabstimmung oder spezielle Fähigkeiten erforderlich wären. BigQuery, die revolutionäre Google-Technologie zur Analyse sehr großer Datenmengen, basiert auf Dremel und bietet eine Performance, für die große Unternehmen bisher Millionen zahlen mussten, zu einem Preis von wenigen Cent pro Gigabyte.

BigQuery ist ein Data Warehouse, das sich optimal für die Ausführung von SQL-Abfragen für umfangreiche, strukturierte und semistrukturierte Datensätze eignet. Beispiele für Anwendungsfälle und Datensätze sind:

- Ad-hoc-Analytics
- Webprotokolle
- Computer-/Serverprotokolle
- Datenbestände aus dem „Internet der Dinge“
- Kundenverhalten im E-Commerce
- Daten mobiler Apps
- Analysen im Einzelhandel
- Telemetriedaten aus Spielen
- Daten aus Google Analytics Premium
- Alle Datensätze, bei denen ein herkömmliches relationales Datenbanksystem (RDBMS) Minuten (oder gar Stunden) braucht, um eine Batch-Abfrage zu verarbeiten

BigQuery ist komplett automatisiert, wartungsfrei und in die Google Cloud Platform integriert. Im Gegensatz zu anderen cloudbasierten Analyselösungen ist für die Nutzung von BigQuery keine Vorabbereitstellung eines Server-Clusters erforderlich. Die Verarbeitungscluster werden von BigQuery während der Ausführung richtig dimensioniert und bereitgestellt.

Wenn Ihre Datenmengen ansteigen, fügt BigQuery automatisch Rechenleistung hinzu – und trotzdem zahlen Sie den gleichen Preis pro Gigabyte.

Legacy SQL im Vergleich zu Standard-SQL

Google BigQuery hat seine APIs einem Upgrade unterzogen und nutzt jetzt Standard-SQL neben BigQuery SQL (das jetzt als Legacy SQL bezeichnet wird), und Tableau hat den Google BigQuery-Konnektor aktualisiert, um diesen Wechsel zu Standard-SQL zu unterstützen. Standard-SQL bietet den BigQuery-Benutzern unter anderem folgende Vorteile: LOD-Ausdrücke, schnellere Metadatenprüfung und die Fähigkeit, mit Ihrer Verbindung ein Abrechnungsprojekt auszuwählen. Diese Anleitung setzt die Nutzung von Standard-SQL voraus.

Weitere Informationen zur Migration von Legacy SQL auf Standard-SQL finden Sie im [Onlinehilfe-Leitfaden](#) unter „[Migrieren von Legacy SQL](#)“.

Tableau

Tableau hilft Benutzern dabei, ihre Daten sichtbar und verständlich zu machen. Unsere moderne Analytics-Plattform basiert auf einer Technologie, die an der Universität Stanford entwickelt wurde. Sie gibt normalen Business-Anwendern das volle Leistungsspektrum von Daten an die Hand. Dadurch kann sich ein breiter Nutzerkreis mit seinen Daten beschäftigen, Fragen stellen, Probleme lösen, Einblicke teilen und einen transformierenden Mehrwert schaffen. Ganz gleich, ob Benutzer mit herkömmlichen BI-Tools vertraut sind, sie lernen schnell, Tableau zum Kreieren und Erkunden von umfangreichen, interaktiven Visualisierungen und leistungsstarken Dashboards mit einer intuitiven Drag-and-Drop-Benutzeroberfläche zu nutzen.

Wir haben die Funktionen der Tableau-Plattform erst kürzlich um visuelle, direkte und intelligente **Datenvorbereitungsfunktionen** und die Fähigkeit, **veröffentlichte Datenquelle mit natürlicher Sprache abzurufen, erweitert**.

Native Optimierungen von Tableau

Datenquellen-Konnektor: Tableau verfügt über einen nativen, optimierten Konnektor für Google BigQuery, der sowohl Live-Daten-Verbindungen als auch In-Memory-Extrakte unterstützt. Mittels Data Blending können Anwender BigQuery-Daten mit Daten aus unseren 67 weiteren unterstützten Datenquellen verknüpfen. Für Visualisierungen, die mit Tableau Server oder Tableau Online in der Cloud veröffentlicht werden, bleibt die direkte Verbindung zu Google BigQuery erhalten.

Parallelabfragen: Tableau kann mithilfe der Möglichkeiten von Google BigQuery und anderen Datenquellen bis zu 16 gleichzeitige Abfragen durchführen. Batches unabhängiger und deduplizierter Anfragen werden in Gruppen zusammengefasst und an BigQuery gesendet, wenn sich das Ergebnis nicht bereits im Cache befindet. Dank der horizontalskalierenden Architektur von BigQuery können Nutzer mit deutlichen Leistungssteigerungen durch Parallelabfragen rechnen.

Query Fusion: Tableau wird, wenn möglich, mehrere Abfragen aus Arbeitsmappen und Dashboards miteinander verschmelzen und dadurch die Anzahl der an BigQuery gesendeten Abfragen reduzieren. Zuerst identifiziert Tableau ähnliche Abfragen, wobei Unterschiede in den zurückgegebenen Spalten ausgeklammert werden. Dann werden Abfragen miteinander verknüpft, bei denen sich nur die Aggregationsebene oder eine Benutzerberechnung unterscheiden.

Externer Abfrage-Cache: Wenn sich die zugrunde liegende Datenquelle seit dem Zeitpunkt der letzten identischen Abfrage das letzte Mal nicht geändert hat, liest Tableau automatisch den zuvor gespeicherten Abfrage-Cache. Dadurch sind nahezu verzögerungsfreie Antwortzeiten möglich.

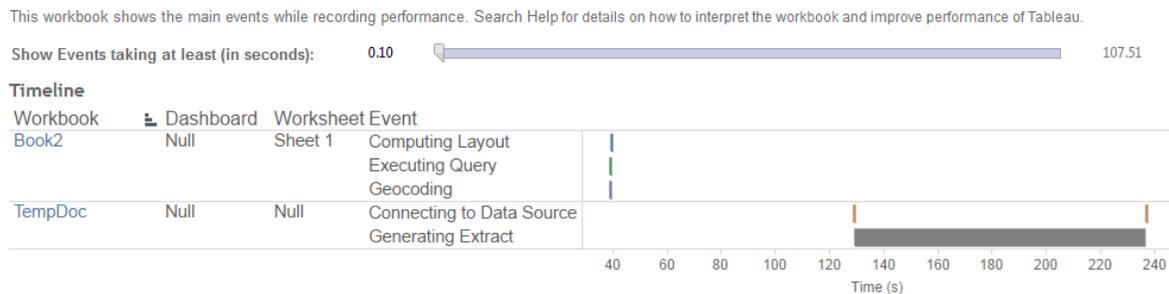
On-Demand-Verbindungen in Tableau Desktop: Wenn Sie eine veröffentlichte Arbeitsmappe öffnen, verbindet sich Tableau Desktop nur mit der Datenquelle, die zum Anzeigen der Daten der aktuellen Tabelle erforderlich ist. Mit anderen Worten: Sie sehen Ihre Daten schneller.

Best Practices für Performance: Tableau

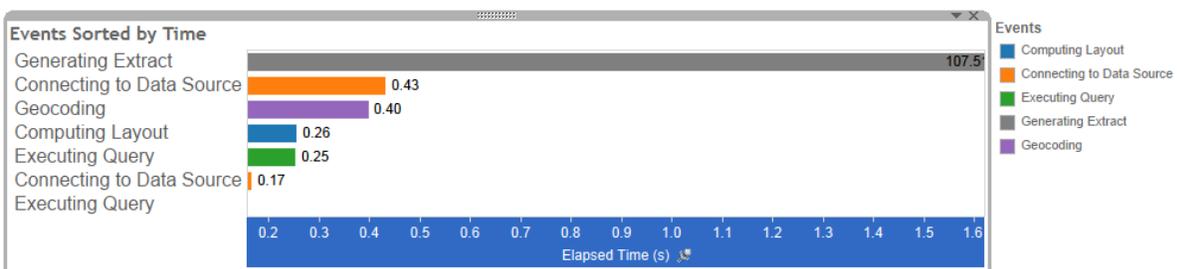
Bevor wir uns weitere Tools und benutzerdefinierte Einstellungen ansehen, empfehlen wir zunächst, nach Möglichkeit mit der Aktualisierung der Tableau-Bereitstellung fortzufahren. Auf diese Weise können Sie von den Leistungsverbesserungen profitieren, die wir fortwährend mit den neuesten Produktversionen einführen.

Leistungsaufzeichnung

Die Leistungsaufzeichnung ist ein integriertes Tool, mit dem Sie langsame Abfragen lokalisieren und Ihre Arbeitsmappen für maximale Leistung optimieren können. Es erfasst für jede einzelne Arbeitsmappe die Zeit, die für das Ausführen einer Abfrage und die Berechnung des Layouts benötigt wird. Wenn der Anwender mit der Maus auf einen der grünen Balken darunter zeigt, wird die entsprechende in BigQuery generierte Abfrage angezeigt. Nachdem Sie eine langsame Abfrage ermittelt haben, können Sie das Leistungsproblem häufig beheben, indem Sie Ihr Datenmodell überarbeiten.



Identifizieren Sie in den Spalten Zeitlinienansicht, Arbeitsmappe, Dashboard und Arbeitsblatt den Kontext für Ereignisse.



Ereignisse von längerer Dauer können Ihnen helfen zu identifizieren, was Sie zur Beschleunigung Ihrer Arbeitsmappe zuerst angehen sollten.

Weitere Informationen zum Erstellen oder Interpretieren einer Leistungsaufzeichnung finden Sie hier:

[Leistungsaufzeichnung in Tableau Desktop \(Erstellen\)](#)

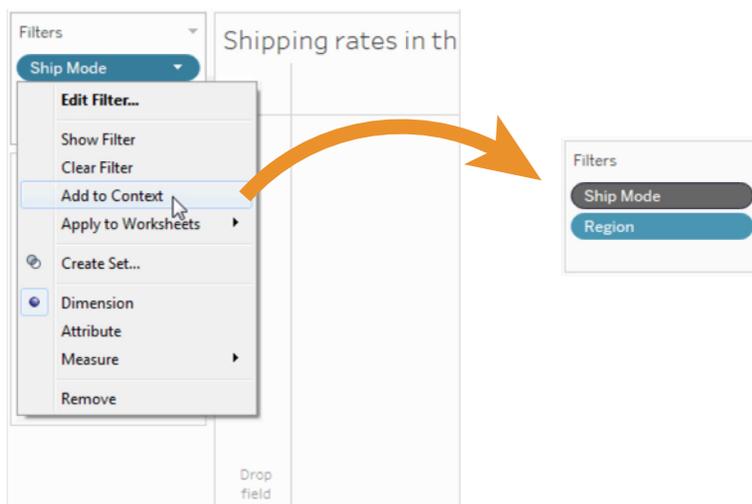
[Leistungsaufzeichnung in Tableau Server \(Interpretieren\)](#)

Kontextfilter

Wenn Sie Filter auf eine große Datenquelle anwenden, können Sie die Leistung verbessern, indem Sie Kontextfilter einrichten. Zuerst wird ein Kontextfilter auf die Datenquelle angewendet, sodass zusätzliche Filter nur auf die resultierenden Datensätze angewendet werden. Diese Reihenfolge verhindert das Anwenden jedes Filters auf jeden Datensatz in der Datenquelle.

Wenn Sie Filter festlegen, welche die Größe des Datensatzes deutlich reduzieren, und diese Filter für mehrere Datenansichten verwenden, sollten Sie sie als Kontextfilter einrichten.

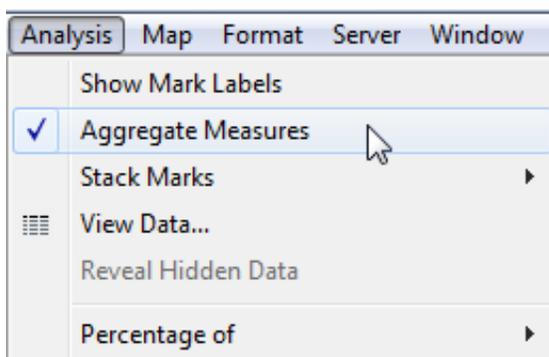
Weitere Informationen zum [Optimieren der Viewsleistung mithilfe von Kontextfiltern](#) finden Sie in unserem [Onlinehilfe-Leitfaden](#).



Sie können zur Leistungssteigerung einen oder mehrere Kontextfilter festlegen.

Aggregieren von Kennzahlen

Falls die von Ihnen erstellten Ansichten langsam sind, stellen Sie sicher, dass Sie mit aggregierten und nicht mit aufgeschlüsselten Kennzahlen arbeiten. Wenn eine Ansicht langsam ist, bedeutet das normalerweise, dass Sie versuchen, zu viele Datenzeilen auf einmal anzuzeigen. Sie können die Anzahl der Zeilen durch Aggregation der Daten reduzieren.



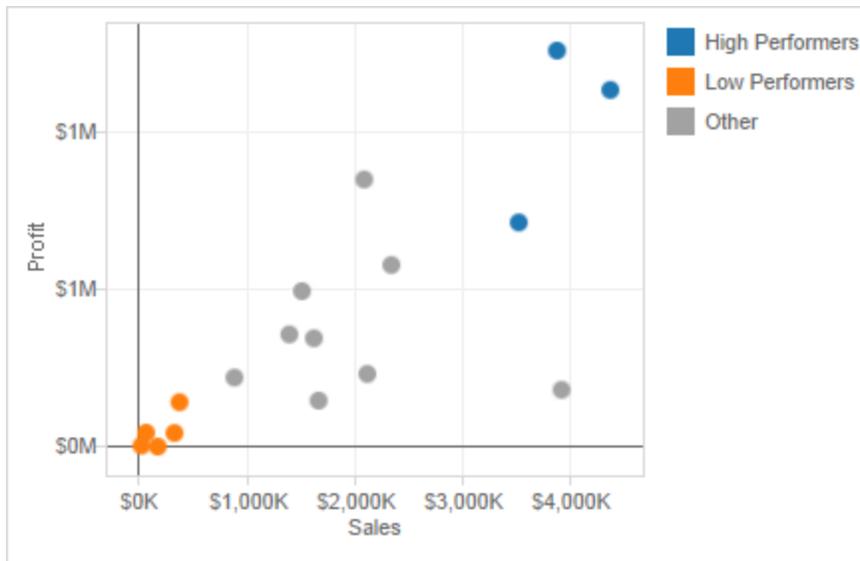
Stellen Sie fest, ob Kennzahlen im Analysemenü aggregiert sind. Sie können für Kennzahlen auch Standardaggregationen festlegen.

Weitere Informationen finden Sie in unserem [Onlinehilfe-Leitfaden zu Datenaggregation](#).

Sets und Gruppen

Wenn Sie eine Dimension filtern möchten, um Elemente anhand einer Reihe von Kennzahlwerten zu entfernen, sollten Sie einen Satz erstellen, anstatt einen quantitativen Filter zu verwenden. Sie können beispielsweise einen Satz erstellen, der anstatt aller Elemente in einer Dimension nur die ersten 50 Elemente ausgibt.

Wenn Sie eine Gruppe aus einer Auswahl erstellen, müssen Sie darauf achten, dass Sie nur die gewünschten Spalten einbeziehen. Jede im Satz enthaltene zusätzliche Spalte beeinträchtigt die Leistung.



Wenn Sie in Tableau Gruppen erstellen, haben Sie die Möglichkeit, alle verbleibenden Elemente in einer Gruppe „Sonstige“ zusammenzufassen.

Weitere Informationen finden Sie in unserem [Onlinehilfe-Leitfaden zum Erstellen von Sätzen](#) und unserem [Onlinehilfe-Leitfaden zum Erstellen von Gruppen](#).

Filter zuerst hinzufügen

Wenn Sie mit einer großen Datenquelle arbeiten und automatische Aktualisierungen deaktiviert haben, kann es passieren, dass eine sehr langsame Abfrage erstellt wird, wenn Sie der Ansicht Filter hinzufügen. Statt die Ansicht aufzubauen und dann Filter festzulegen, sollten Sie erst die Filter festlegen und dann Felder in die Ansicht verschieben. Auf diese Weise werden die Filter zuerst angewendet, wenn Sie die Aktualisierung vornehmen oder automatische Aktualisierungen wieder aktivieren.

Deaktivieren der automatischen Aktualisierungen

Wenn Sie ein Feld auf einen Container ziehen, erstellt Tableau die Ansicht durch automatische Abfrage der Datenquelle. Wenn Sie eine komplexe Datenansicht erstellen, können Abfragen zeitaufwändig sein und die Systemleistung deutlich beeinträchtigen. In diesem Fall können Sie Tableau anweisen, Anfragen zu deaktivieren, während Sie die Ansicht erstellen. Sie können dann die Abfragen wieder aktivieren, sobald Sie zur Anzeige der Ergebnisse bereit sind.

Weitere Informationen finden Sie in unserem [Onlinehilfe-Leitfaden zu automatischen Aktualisierungen und Leistung](#).

Auf Warnungen achten

Tableau zeigt ein Dialogfeld mit einer Leistungswarnung an, wenn Sie versuchen, eine große Dimension (mit vielen Elementen) auf einen Container zu ziehen. Im Dialogfeld stehen, wie in der Abbildung unten zu sehen, vier Optionen zur Auswahl. Wenn Sie sich dafür entscheiden, alle Elemente hinzuzufügen, müssen Sie mit deutlichen Leistungseinbußen rechnen.

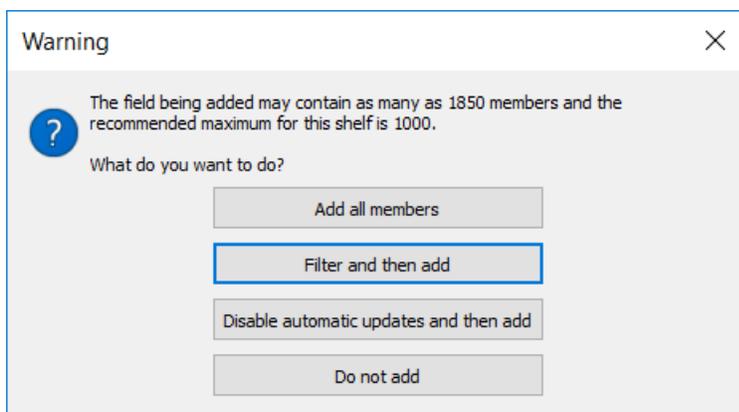


Tableau warnt Sie, wenn durch die Platzierung einer großen Dimension in einen Container Leistungseinbußen zu erwarten sind.

Optimieren paralleler Abfragen

Sie können Attribute zur individuellen Anpassung nutzen, um so durch die Konfiguration paralleler Abfragen die Leistung von Ergebnissets, die von BigQuery nach Tableau Online und Tableau Server sowie auf Tableau Desktop ausgegeben werden, zu verbessern. Diese Attribute zur individuellen Anpassung können in Ihre veröffentlichte Arbeitsmappe oder Ihre Datenquelle übernommen werden, wenn Sie die Attribute festlegen, bevor Sie die Arbeitsmappe oder die Datenquelle in Tableau Online oder Tableau Server veröffentlichen.

Weitere Informationen finden Sie in unserem [Onlinehilfe-Leitfaden für Google BigQuery](#) unter „Attribute zur individuellen Anpassung zur Verbesserung der Abfrageleistung verwenden“.

Best Practices für Kosten und Leistung: Google BigQuery

Für Hochleistungsabfragen und zur Reduzierung der Kosten ist es prinzipiell am besten, keine Verbundtabellen zu verwenden, deren Daten aus einer externen Datenquelle wie Google Cloud Storage stammen. Wenn Sie in einem solchen Fall iterative Abfragen des Datensatzes durchführen möchten, sollten Sie die Abfrage-API benutzen, um die Daten – unabhängig von Tableau – in BigQuery zu materialisieren, um Hochleistungsabfragen des Datensatzes mit Tableau zu ermöglichen

Denormalisierung und Vorabverknüpfung

BigQuery unterstützt sehr große JOINS (Verknüpfungen) und die JOIN-Leistung ist sehr gut. Davon abgesehen ist BigQuery ein spaltenorientierter Datastore und maximale Leistung wird mit denormalisierten Datensätzen erreicht.

Ein Vorteil der Cloud ist die Fähigkeit, Speicher- und Rechenressourcen zu entkoppeln und so den Benutzern das Skalieren und die individuelle Bezahlung für jede Ressource zu ermöglichen. Da BigQuery-Speicherplatz sehr preisgünstig und skalierbar ist, empfiehlt es sich oftmals, Datensätze zu denormalisieren und vorab zu verknüpfen, damit homogene Tabellen entstehen. Dies bedeutet im Wesentlichen, dass Sie weniger Rechenressourcen, dafür aber mehr Speicherressourcen nutzen werden (letztere sind leistungsstärker und kostengünstiger). Da BigQuery ein spaltenbasierter Speicher ist und Daten gut komprimieren kann, ist die Reduzierung der Rechenressourcen zugunsten von mehr Speicherplatz keine schlechte Wahl und vermutlich auch kostengünstiger.

BigQuery ist ein hervorragendes ETL-Tool, mit dem Sie umfangreiche Transformationen und Pipelines schnell und effizient ausführen können. Achten Sie darauf, dass „Allow Large Results“ bei der Materialisierung von Datensätzen mit mehr als 128 MB aktiviert ist.

Weitere Informationen zur Vorbereitung der Daten für das Laden und zum Abfragen von Daten mithilfe des SQL-Dialekts in BigQuery finden Sie hier:

[Laden von denormalisierten, verschachtelten und wiederholten Daten](#)
[Schreiben umfangreicher Abfrageergebnisse](#)

Horizontale Fragmentierung von Tabellen nach Datum

Das Unterteilen einer Tabelle in kleinere Partitionen – die so genannte horizontale Fragmentierung – kann helfen, das Datenmanagement zu vereinfachen, die Abfrageleistung zu verbessern und die Kosten zu senken. Darüber hinaus unterstützt BigQuery das Clustering über eine partitionierte Tabelle. Dies ist hilfreich, wenn Daten bereits nach einer Datums- oder Zeitstempelspalte partitioniert sind oder wenn Sie Filter oder eine Aggregation auf bestimmte Spalten in Ihren Abfragen angewendet haben.

Einige Daten eignen sich prinzipiell für eine Fragmentierung nach dem Datum: z. B. Protokolldaten oder sämtliche Daten, bei denen die Datensätze einen monoton ansteigenden Zeitstempel enthalten. Fragmentieren Sie Ihre BigQuery-Tabellen in diesem Fall nach Datum und fügen Sie das Datum in den Tabellennamen ein. Um hiervon zu profitieren, müssten Sie in Tableau benutzerdefinierte SQL-Abfragen nutzen.

Weitere Informationen finden Sie in [unserem Onlinehilfe-Leitfaden zur Herstellung einer Verbindung zu einer benutzerdefinierten SQL-Abfrage](#).

Benennen Sie beispielsweise Ihre Tabellen wie folgt: meineTabelle__20170501, meineTabelle__20170502 usw. Wenn Sie dann eine Abfrage ausführen, die nach Datum filtert, verwenden Sie die BigQuery-Funktion „Wildcard Table“ (Platzhaltertabelle):

```
SELECT
  name
FROM
  `myProject.myDataSet.mytable_*`
WHERE
  age >= 35
```

Das obige Beispiel schließt automatisch alle Tabellen mit dem Präfix „mytable_“ ein.

Um Platzhalter verwenden zu können, müssen Sie Ihre Tabellen gemäß dem folgenden Muster benannt haben: [arbitrary prefix]YYYYMMDD.

Andere Datenbanksysteme nutzen die horizontale Fragmentierung zur Verbesserung der Leistung. Der bei BigQuery durch horizontale Fragmentierung nach Datum erzielte Leistungsunterschied kann vernachlässigt werden; der Hauptgrund sind hier die Kosten. Da Sie weniger Daten verarbeiten, bezahlen Sie auch weniger Geld pro Abfrage.

Hier gilt allerdings folgender Vorbehalt: Falls Sie beschließen sollten, auf sehr kleiner Ebene zu fragmentieren, könnten Sie zu viele Fragmente erhalten, was sich direkt auf die Leistung auswirkt. Sie müssen darauf achten, dass Sie nicht zu stark fragmentieren. Alles oberhalb von „täglich“ ist im Allgemeinen akzeptabel.

Weitere Informationen zur horizontalen Fragmentierung finden Sie hier:

[Einführung in partitionierte Tabellen](#)

[Einführung in geclusterte Tabellen](#)

[Abfragen mehrerer Tabellen mithilfe einer Platzhaltertabelle](#)

Angabe einer Zieltabelle beim Ausführen vieler ähnlicher Abfragen

Obwohl die Abfrage-Zwischenspeicherung hilfreich ist, wenn Sie viele identische Abfragen durchführen, nützt sie bei ähnlichen, aber geringfügig unterschiedlichen Abfragen gar nichts (wenn z. B. nur die Werte in einer WHERE-Klausel zwischen den Abfragen geändert werden). Führen Sie in diesem Fall eine Abfrage für Ihre Quellentabelle durch und schreiben Sie die Datensätze, die Sie wiederholt abfragen werden, in eine neue Zieltabelle. Führen Sie dann Abfragen für die neu erstellte Zieltabelle durch.

Nehmen wir zum Beispiel an, Sie möchten drei Abfragen mit drei unterschiedlichen WHERE-Klauseln ausführen:

```
WHERE col1 = "a"
```

```
WHERE col1 = "b"
```

```
WHERE col1 = "c"
```

Führen Sie eine Abfrage für Ihre Quellentabelle aus und schreiben Sie die Ausgabedatensätze in eine Zieltabelle:

```
SELECT col1
```

```
FROM source
```

```
WHERE col1 = "a" OR col1 = "b" OR col1 = "c"
```

Indem wir die WHERE-Klauseln mit OR verbinden, erfassen wir alle relevanten Datensätze. Unsere neue Zieltabelle ist potenziell viel kleiner als die ursprüngliche Quellentabelle. Da die Kosten für BigQuery auf der in einer Abfrage verarbeiteten Datenmenge basieren, ist es günstiger nachfolgende Abfragen für die neue Zieltabelle auszuführen als Abfragen direkt für die Quelldatei. Es muss darauf geachtet werden, diese Tabellen künftig zu bereinigen, damit sich die Speicherkosten für diese Tabellen nicht akkumulieren.

Google BigQuery ML-Modellresultate mit Tableau visualisieren

Mit BigQuery ML können Benutzer Modelle anhand der in BigQuery gespeicherten Daten mithilfe eingebetteter Machine Learning-Technologie „anlernen“. Wie jedoch bei der Arbeit mit anderen Daten auch, ist die direkte Abfrage einer Datenbank nicht immer die ideale Methode zum Erkunden der Ergebnisse Ihres Modells.

Mit seinem BigQuery-Konnektor ermöglicht Ihnen Tableau die einfache Bearbeitung der Ergebnisse Ihres Prognosemodells in einer Art, die ein intuitives Verständnis der Daten vereinfacht. Darüber hinaus haben Benutzer mit Tableau eine einfache Möglichkeit zum Teilen ihres Modells und der Modellergebnisse mit anderen, sodass diese von all Ihrer Arbeit profitieren können.

Ein Beispiel finden Sie unter [Verwenden von BigQuery ML mit Tableau zur Prognose von Wohnungspreisen](#).

Zwar ist eine benutzerdefinierte SQL in dieser Übung erforderlich, um einen Machine Learning-Algorithmus in Google BigQuery ML aufzurufen, bei der Anwendung einer benutzerdefinierten SQL statt den nativen Verbindungen von Tableau in anderen Szenarien sollte jedoch stets die Leistung berücksichtigt werden. Nutzen Sie für eine optimierte Leistung nach Möglichkeit die nativen Datenquellenverbindungen von Tableau.

Weitere Informationen zum Erstellen und Ausführen von Machine Learning-Modellen mithilfe von SQL-Standardabfragen finden Sie im [Onlinehilfe-Leitfaden zu Google Big Query ML](#).

Fallstudie: Top-Tipps für Selfservice-Analytics mit Tableau und Google BigQuery von zulily

zulily ist ein schnell wachsendes E-Commerce-Unternehmen, das unter Verwendung von Google BigQuery als Business-Data-Warehouse und Tableau für Datenzugriff und visuelle Analytics eine Big Data-Plattform erstellt hat. Zum einen kann die Analytics durch die Integration von BigQuery und Tableau im Handumdrehen Daten abrufen, verarbeiten und basierend darauf Berichte und Modelle erstellen, ohne für alltägliche Aufgaben die IT-Abteilung hinzuziehen zu müssen. Zum anderen haben Geschäftsanwender in Echtzeit Zugriff auf wichtige Daten, die als Grundlage für schnelle Entscheidungsfindungsprozesse dienen, und können einfache Erkenntnisse ohne Unterstützung durch Analysten selbst gewinnen.

Hier sind einige der Best Practices von zulily:

Reduzieren der Latenz durch die Verwendung von Tableau Server auf der Google-Compute Engine: Statt einem herkömmlichen Modell, bei dem Regionen separate VPCs sind, können Sie den privaten Google-Backbone nutzen, ohne dabei eine Internetverbindung herzustellen – und ohne zusätzliches Setup. Dadurch können Sie Ihre Bereitstellungen korrekt dimensionieren, ohne sie zu groß anzulegen.

Verwendung von föderierten Quellen und Verweisen von Tableau auf BigQuery: Für Daten in der Google Cloud sollten Sie sich die Fähigkeit von BigQuery zunutze machen, externe Datenquellen abzufragen. Sie sollten BigQuery als Ihren Data Lake nutzen. In bestimmten Szenarien reduzieren Sie die Datenmenge, die zu Analysezwecke über das Netzwerk und in Tableau verschoben werden muss.

Verarbeiten großer Datensätze mit einer Direktverbindung in BigQuery: Nutzen Sie die Fähigkeit von BigQuery zur Verarbeitung großer Datensätze sowie die Fähigkeit, nur Ergebnisse über das Netzwerk zu transferieren. Legen Sie Ihre Tableau-Standardverbindung zu BigQuery als Direktverbindung fest, sofern Sie keinen bestimmten Grund für die Datenextraktion haben.

Weitere Informationen sowie die komplette Liste aller 10 Tipps erhalten Sie in der zweiteiligen Blogreihe:

Teil 1: [Warum zulily eine Plattform für Selfservice-Marketing-Analytics mit Tableau und Google BigQuery eingerichtet hat](#)

Teil 2: [Die Top 10 der Tipps von zulily für Selfservice-Analytics mit Google BigQuery und Tableau](#)

Fazit

Durch die Anwendung von Best Practices können Geschäftsanwender und Datenanalysten die Leistung und Reaktionsfähigkeit von Tableau-Visualisierungen, die für Google BigQuery erstellt wurden, maximieren. Wenn diese Technologien kombiniert werden, können die Anwender Milliarden von Datenzeilen blitzschnell visualisieren.

Über Tableau

Tableau ist eine vollumfängliche, anwenderfreundliche, unternehmensfähige visuelle BI-Plattform, die den Benutzern hilft, ihre Daten zu sehen und verstehen – und zwar mithilfe von blitzschnellen Selfservice-Analytics im großen Maßstab. Egal ob Tableau auf lokalen Computern oder in der Cloud bzw. in einer Windows- oder Linux-Umgebung ausgeführt wird, die Lösung nutzt stets Ihre vorhandenen Technologie-Investitionen und lässt sich mühelos skalieren, wenn Ihre Datenumgebung verlagert wird und wächst. Optimale Nutzung Ihrer wertvollsten Ressourcen – Ihre Daten und Ihre Mitarbeiter

Weitere Materialien

[Kostenlose Tableau-Testversion](#)

[Tableau Onlinehilfe-Leitfaden: Google BigQuery](#)

[Tableau Server und Google Cloud Platform: blitzschnelle Business Intelligence in der Cloud](#)

[Tableau- und Google-Lösungen](#)

[Tableau und Big Data: eine Übersicht](#)

[Warum Business-Analytics in der Cloud?](#)

[Entwerfen effizienter Arbeitsmappen](#)

