

# Tableau und Big Data: Ein Überblick



# Inhoudsopgave

<b>Hoe big data er tegenwoordig uitziet</b> .....	<b>3</b>
De evolutie van data en de vraag naar analyses .....	3
Big data is zowel een vloek als een zegen.....	4
<b>Hoe Tableau werkt met big data</b> .....	<b>5</b>
De 'big (data) picture'.....	5
Datatoegang en -connectiviteit .....	5
Snelle interactie met alle data op niveau.....	6
<b>Tableau en het big data-analyse-ecosysteem</b> .....	<b>7</b>
Cloudinfrastructuur .....	8
Opname en voorbereiding .....	8
Opslag en verwerking.....	9
Queryversnelling .....	10
Datacatalogus.....	10
<b>Big data-analysearchitecturen</b> .....	<b>10</b>
Voorbeelden van grote cloudproviders.....	11
Voorbeelden van Tableau-klanten .....	12
Veelvoorkomende patronen .....	13
<b>Over Tableau en extra bronnen</b> .....	<b>14</b>



# Hoe big data er tegenwoordig uitziet

## De evolutie van data en de vraag naar analyses

Data is overall, net als de vraag om deze in te zien en te analyseren. 'Big data' raakt als buzzwoord al steeds meer ingeburgerd, maar de 'drie V's' van big data (volume, variety (variëteit) en velocity (snelheid)) zijn als nooit tevoren van toepassing op big data-analyse. Deze en andere V's die in de branche aan de orde komen (zoals variabiliteit, geldigheid, juistheid etc.) geven aan dat big data eigenlijk ook gewoon data is. Alleen is het allemaal zo complex geworden, dat organisaties innovatieve manieren moeten ontwikkelen om het effectief te verzamelen, onderhouden en gebruiken.

Er vindt een digitale transformatie plaats in elke branche en in alle soorten organisaties. Hierbij wordt door een meervoud van 'dingen' een enorme hoeveelheid data aangemaakt; dit in vele indelingen en uit vele bronnen. Organisaties verzamelen, verwerken en analyseren meer diverse data dan ooit tevoren. Van schemavrije JSON's tot geneste typen in andere databases zoals relationele en NoSQL, tot non-flat data (bijv. Avro, Parquet, XML etc.): er komen steeds meer data-indelingen en connectors zijn cruciaal om ze te kunnen gebruiken.

### Organisaties beschikken vaak over een combinatie van het volgende:

- **Gestructureerde data** met voorberekende aggregaten voor specifieke vragen, mogelijk opgehaald als extracten voor in-memory-berekening en geaggregeerd voor analyse. Dit is meestal de meest verfijnde en gemakkelijk toegankelijke data van een organisatie.
- **Deels gestructureerde data** (of objectopslag) mogelijk in relationele databases, datawarehouses en datamarts. Vaak zijn dit zakelijke concepten voor entiteitsanalyse (bekende vragen met onbekende antwoorden) die regelmatig worden vernieuwd. Denk hierbij aan transacties, mogelijkheden of acties die worden opgepikt door afzonderlijke verkopers naar aanleiding van mogelijkheden.
- **Onbewerkte, ongestructureerde data** in een data lake of cloudopslag. Dit omvat onder andere stroomdata van sociale-netwerkfeeds en IoT-apparaten. Datawetenschappers kunnen deze data verzamelen en transformeren, maar het volledige potentieel is nog steeds onbekend.

Hoewel voor sommige data de waardevolste gebruikstoepassingen nog moeten worden gezocht, is er voor al deze data een steeds grotere vraag naar kenniswerkers die het ten behoeve van besluitvorming kunnen analyseren en inzetten. De gebruikte toepassingen voor data-analyse en visualisatie komen steeds dichter rondom de data zelf te liggen. Dit betekent een grootschalige verschuiving naar de cloud, waar analyses naast robuuste opslag en dataverwerking komen te liggen voor een grotere flexibiliteit en schaal. Of een organisatie nu beschikt over een uitgebreid big data-systeem in de cloud of momenteel slechts weinig analyses verricht op haar data, er kan aanzienlijke vooruitgang worden geboekt als de mensen binnen het bedrijf en de IT-afdelingen de mogelijkheid wordt geboden om patronen te visualiseren en te analyseren om beter inzicht te verkrijgen.



Ondanks dat moderne analysemethoden ruimere mogelijkheden scheppen voor meer zakelijke gebruikers op alle vaardigheidsniveaus, vormt het vinden van manieren om van al deze data een geschikte bron te maken voor de volledige organisatie een complexe uitdaging. De behoeften van bedrijven veranderen zo vaak als de data zelf, waardoor een big data-strategie en -architectuur noodzakelijk zijn om wendbaar en flexibel te zijn. In plaats van enorme platformen met een focus op dataconnectiviteit bouwen, kunnen organisaties beter hun kijk op big data-mogelijkheden verruimen en nadenken over de ontwikkeling van gebruikstoepassingen voor analyse. Anders zouden ze weleens het grotere plaatje uit het oog kunnen verliezen.

## **Big data is zowel een vloek als een zegen**

Data assets vormen steeds vaker een belangrijk onderscheidingsfactor tussen extreem winstgevende en noodlijdende bedrijven. Maar de enorme schaal, groei en diversiteit van data is eenvoudigweg te omvangrijk en te duur om in combinatie met relationele-databasebeheersystemen te kunnen gebruiken. Naast een besparing op de hardwarekosten middels voor- en gedeelde berekeningen, zoeken klanten ook naar een werkmethode waarbij de data zo weinig mogelijk hoeft te worden verplaatst. Infrastructuur waarmee data op de meest flexibele wijze kunnen worden verplaatst, helpt ook met het aanpakken van het gat tussen onbewerkte, ongestructureerde data en data die gebruikers meteen kunnen analyseren.

Organisaties lopen ook tegen problemen met connectiviteit en prestaties aan, maar zelfs met opties voor live-verbindingen en in-memory analyses kunnen enorme data lakes een zware druk leggen op bewerkingen voor het genereren van extracten of het combineren van andere data. Een moderne en selfservice-aanpak ten opzichte van analytics biedt een grote flexibiliteit, maar door het leggen van een enorm aantal joins kan het systeem dichtslibben.

IT en bedrijven moeten samenwerken, maar met een omgekeerde methodologie bestaande uit onderwerpexperts die metadata, bedrijfsregels en rapportagemodellen opstellen. Deze processen moeten voortdurend worden bijgesteld en verbeterd om te kunnen voldoen aan de veranderende industrie; in het hedendaagse tijdperk van digitale transformatie staan bedrijven niet stil, dus jouw data-analyseraamwerk zou dat ook niet moeten doen.




# Hoe Tableau werkt met big data

## De 'big (data) picture'

Met alles wat we doen bij Tableau, ondersteunen we onze missie om mensen hun data beter te leren inzien en begrijpen. Tableau is het moderne analyseplatform voor de digitale economie omdat we fundamenteel geloven in de democratisering van data. De mensen die data het beste kennen, moeten de mogelijkheid hebben om vragen te stellen aan de data. Kenniswerkers op alle vaardigheidsniveaus moeten inzichten in hun data kunnen gebruiken, analyseren en ontdekken, waar die dan ook te vinden zijn.

Omdat veel klanten te maken hebben met een verscheidenheid aan big data-technologieën, hebben we onze investeringen op het gebied van engineering, samenwerkingen binnen het ecosysteem en algehele visie afgestemd op de evolutie van het datalandschap. Tableau heeft een rijke historie van investeringen die in de aanloop naar de toename in big data zijn gedaan. Deze investeringen omvatten dataconnectiviteit met zowel Hadoop- als NoSQL-platformen, alsmede grootschalige on-premises en clouddatawarehouses.

 We zijn ooit begonnen met een compacte zakelijke gebruikstoepassing die zich snel uitbreidde. Big data-analyse leeft bij iedereen, en Tableau maakt het toegankelijk.

—ASHISH BRAGANZA, DIRECTOR OF GLOBAL BUSINESS INTELLIGENCE BIJ LENOVO

[Ontdek hoe Lenovo rapportering in ruim 28 landen 95% efficiënter heeft gemaakt](#)

## Datatoegang en -connectiviteit

Om data-analyses van verschillende grootte of in andere indelingen te kunnen gebruiken, ondersteunen we een brede toegang tot data, waar deze zich ook bevinden. Tableau ondersteunt nu meer dan 75 oorspronkelijke dataconnectors, alsmede talloze andere door middel van onze uitbreidingsmogelijkheden. Zolang er nieuwe databronnen beschikbaar komen en van belang worden voor onze gebruikers, blijven we de connectors van leveranciers integreren en certificeren met Tableau, zodat het gemakkelijker wordt om toegang te krijgen tot data. Wij zijn van mening dat er altijd een heleboel databronnen zijn of zullen zijn die iemand graag wil gebruiken, of het nu gaat om webverkeer, records in databases, logbestanden of andere zaken.

- **SQL-verbindingen** — Tableau maakt gebruik van SQL om gegevens uit te wisselen met Hadoop, NoSQL-databases en Spark. De SQL die Tableau genereert, is afgestemd op de ANSI SQL-92-standaard. SQL is een krachtig middel omdat het uitermate compact (één expressie), open source en gestandaardiseerd is. Het is niet afhankelijk van een bibliotheek en bijzonder rijk en expressief. Met SQL kan men bijvoorbeeld join operations, functies, criteria, samenvattingen, gegroepeerde en geneste bewerkingen uitdrukken.



- **NoSQL-interfaces** — NoSQL (Not only SQL)-databases kunnen, precies zoals de naam al zegt, data gebruiken die naast in relationele ook in niet-relationele indelingen is gemodelleerd, waarbij extra opslagtypen worden ondersteund, inclusief kolom, document, sleutelwaarde en grafiek. Tevens betekent dit dat SQL-achtige interfaces worden ondersteund.
- **ODBC** — Tableau gebruikt drivers die de Open Database Connectivity (ODBC)-programmastandaard gebruiken als vertaallaag tussen SQL en SQL-achtige data-interfaces die door deze big data-platformen worden verschaft. Met behulp van ODBC heb je toegang tot alle databronnen die de SQL-standaard ondersteunen en de ODBC API implementeren. Voor Hadoop omvat dit interfaces zoals Hive Query Language (HiveQL), Impala SQL, BigSQL en Spark SQL. Voor de allerbeste prestaties kunnen we de SQL nauwkeurig aanpassen, maar ook aggregaties, filters en andere SQL-bewerkingen genereren en doordrukken naar big data-platformen.
- **Web Data Connector** — Met de Tableau Web Data Connector kunnen SDK-gebruikers connecties maken met data die buiten de bestaande connectors liggen. Gebruikers van selfservice-analyse kunnen hun big data-analyses verbeteren met externe data door een verbinding te maken met bijna alle data die toegankelijk is via HTTP, inclusief interne webservices, JSON-data en REST API.

## Snelle interactie met alle data op niveau

Wij willen dat gebruikers toegang hebben tot al hun data op niveau, zodat deze kunnen worden geïntegreerd met andere data en er snel inzicht kan worden verkregen. Om selfservice visuele analytics met big data mogelijk te maken, heeft Tableau in baanbrekende technologieën geïnvesteerd.

- **Hyper data-engine** — **Hyper** is onze high-performance, in-memory data-enginetechnologie waarmee klanten grote en complexe datasets sneller kunnen analyseren. Door gebruik te maken van een eigen dynamische code en geavanceerde parallele technieken, kan Hyper moderne hardware beter inzetten en zo een 3x betere extractiemogelijkheid en een 5x snellere querysnelheid dan de vorige Tableau-data-engine realiseren. Hyper kan ook tragere databronnen verbeteren en versnellen door een data-extract te maken en dat in het geheugen te laden.
- **Hybride data-architectuur** — Tableau kan live een verbinding maken met databronnen of data (of een subset) in het geheugen laden. Je kunt naar wens schakelen tussen deze modi. Onze hybride aanpak met betrekking tot de toegankelijkheid van data biedt gebruikers een heleboel flexibiliteit om zo optimale queryprestaties te verkrijgen.



- **VizQL™** — Het kloppende hart van Tableau is een zelfontwikkelde technologie waarbij interactieve datavisualisatie een integraal is voor het begrijpen van data. In een traditionele analysetool moet je data analyseren in rijen en kolommen, een subset kiezen van de data die je wilt presenteren, je data omzetten in een tabel en vervolgens een grafiek maken van die tabel. VizQL slaat deze stappen over en maakt direct een visuele weergave van jouw data, waarbij je visuele feedback krijgt tijdens de analyse. Met VizQL kun je onbeperkt je data verkennen om de beste weergave ervan te vinden, en dankzij de onbeperkte mogelijkheden om dingen ongedaan te maken, is er geen verkeerde afslag. In deze cyclus van visuele analyse leren gebruikers terwijl ze bezig zijn, kunnen ze eventueel meer data toevoegen en wordt er uiteindelijk een dieper inzicht gerealiseerd. De ervaring is niet alleen rijker maar ook toegankelijker voor alle vaardigheidsniveaus, iets wat niet kan als dashboards met code moeten worden opgebouwd.

“ Met Tableau kun je echt in real time met de dataset interacteren, en kun je deze analyseren en binnen een paar minuten precies zo presenteren zoals je dat zelf wilt.

— JAMIE FAN, PRODUCT ANALYTICS LEAD BIJ GRAB

[Ontdek hoe je met Grab miljoenen rijen data kunt analyseren en de klantervaring kunt verbeteren](#)

## Tableau en het big data-analyse-ecosysteem

Een modern analyseplatform als Tableau is mogelijk de sleutel waarmee het potentieel van big data en het verkrijgen van inzicht kan worden ontgrendeld, maar het is slechts een deel van de kritische elementen van een complete big data-platformarchitectuur. Een volledige big data-analysepijplijn samenstelling, kan een gigantische uitdaging lijken. Het goede nieuws is, dat je niet een volledig ecosysteem hoeft op te zetten voordat je aan de slag kunt, en ook niet elk afzonderlijk onderdeel hoeft te integreren om een volledige strategie op poten te zetten.

Tableau past keurig in het big data-paradigma omdat we flexibiliteit op de voorgrond hebben gezet: de mogelijkheid om data te verplaatsen tussen platformen, infrastructuur aan te passen wanneer dat nodig is, te profiteren van nieuwe datasoorten en nieuwe gebruikers en gebruikstoepassingen in te zetten. Wij vinden dat het implementeren van een analyse-oplossing voor big data niet hoeft te bepalen hoe je infrastructuur of strategie eruit moet zien, maar dat het je moet helpen de investeringen die je hebt gedaan, te maximaliseren, inclusief die met partnertechnologieën binnen het big data-ecosysteem.



## Cloudinfrastructuur

Organisaties verplaatsen hun bedrijfsprocessen en infrastructuur steeds meer naar de cloud. Omdat de cloudgebaseerde infrastructuur en dataservices een aantal van de belangrijkste struikelblokken met betrekking tot on-premises Hadoop-data lakes hebben weggenomen, zijn cloudgebaseerde big data-analyseoplossingen gemakkelijker te implementeren dan ooit tevoren.

Hadoop heeft de basis gelegd voor het moderne data lake met de krachtige combinatie van goedkope, uitbreidbare opslag (Hadoop Distributed File System—HDFS), speciaal ontworpen verwerkings-engines (eerst MapReduce, en daarna Hive, Impala en Spark) en gedeelde datacatalogus (Hive metastore).

Tegenwoordig kunnen de eenmaal samengevoegde storage- en compute-services naar behoefte en onafhankelijk in de cloud worden geschaald. Bronnen volgen het ophoog en omlaag schalen een stuk eenvoudiger, en bovendien met on-demand-prijzvoering. De cloud zorgt over het algemeen voor een betere efficiency, beter beheer en gemakkelijkere coördinatie van services.

Meer informatie is te vinden in [dit geweldige artikel](#) van Josh Klahr, VP Product bij AtScale.

Tableau levert essentiële integraties met cloudtechnologieën die organisaties al gebruiken, waaronder [Amazon Web Services](#), [Google Cloud Platform](#) en [Microsoft Azure](#).

## Opname en voorbereiding

Bij moderne 'ingest-and-load'-ontwerppatronen is de bestemming voor onbewerkte data in welke omvang of vorm dan ook vaak een data lake: een opslagarchief waarin een enorme hoeveelheid data wordt bewaard in de eigen indeling, ongeacht of dat gestructureerd, deels gestructureerd of ongestructureerd is. Data lakes ondersteunen moderne vereisten voor voor big data-analyse door middel van snellere, flexibelere data-opname en opslag voor iedereen om snel onbewerkte data op verschillende manieren te kunnen analyseren.

Er wordt een voortdurende datastroom gegenereerd door gekoppelde apparaten en apps op verschillende locaties, zoals sociale netwerken, slimme meters, thuisssystemen, videogames en IoT-sensoren. Vaak worden deze data verzameld via pijplijnen of deels gestructureerde data. Hoewel analyses in real time en voorspellende algoritmes op stromen kunnen worden toegepast, worden datastromen meestal in onbewerkte indelingen via een lambda-architectuur voor analysetoepassingen doorgestuurd naar en opgeslagen in een data lake, zoals Hadoop. Een lambda-architectuur is een dataverwerkingsarchitectuur die is ontworpen voor het verwerken van enorme hoeveelheden data door middel van zowel batch- als streamverwerkingsmethoden. Dit ontwerp is een balans van uitdagingen op het gebied van latentie, doorvoer en fouttolerantie. Tegenwoordig is er een grote verscheidenheid aan opties voor het streamen van data en kan men kiezen uit bijvoorbeeld Amazon Kinesis, Storm, Flume, Kafka en Informatica Vibe Data Stream.





Data lakes bieden ook geoptimaliseerde verwerkingsmechanismes via API's of SQL-achtige talen om onbewerkte data te transformeren met 'schema on read'-functionaliteit. Zodra data terecht zijn gekomen in een data lake, moeten deze worden opgenomen en voorbereid voor analyse. Met partners als **Informatica**, **Alteryx**, **Trifacta** en Datameer zorgt tableau ervoor dat dit proces wordt ondersteund en naadloos werkt met Tableau. Voor selfservice-datavoorbereiding kun je **Tableau Prep** gebruiken.

## Opslag en verwerking

Hadoop is gebruikt voor data lakes vanwege de veerkracht en lage kosten, de uitbreidbare dataopslag, de parallelle verwerking en het geclusterde workloadbeheer. Hoewel Hadoop vaak wordt gebruikt als big data-platform, is het geen database. Hadoop is een opensourcesoftwareframework voor het opslaan van data en het uitvoeren van applicaties op clusters van standaardhardware. Het biedt enorme opslagmogelijkheden voor allerlei soorten data, uitstekende verwerkingskracht en de mogelijkheid om gigantische aantallen taken of opdrachten gelijktijdig te verwerken.

In een moderne analyse-architectuur biedt Hadoop goedkope opslag- en data-archiefmogelijkheden voor het offloaden van historische data van het datawarehouse naar online koude opslag. Het wordt ook gebruikt voor IoT-, datawetenschaps- en ongestructureerde-analysegebruikstoepassingen. Tableau biedt rechtstreeks connectiviteit aan alle belangrijke Hadoop-distributies met **Cloudera** via Impala, **Hortonworks** via Hive en **MapR** via Apache Drill.

In een moderne analyse-architectuur is altijd ruimte voor databases en datawarehouses, en blijven deze bovendien een cruciale rol spelen in het leveren van beheerde, nauwkeurige, geconformeerde dimensionale data over de gehele bedrijfsvoering voor selfservice-rapportage. Zelfs bedrijven die overstappen op andere technologieën (bijv. Hadoop, data lakes) behouden meestal relationele databases als onderdeel van de databroncombinatie. **Snowflake** is een voorbeeld van een SQL-gebaseerd datawarehouse voor de bedrijfscloud met een eigen Tableau-connector.

Objectopslag, zoals Amazon Web Services Simple Storage Service (S3) en NoSQL-databases met flexibele planning, kan ook worden gebruikt als data lake. Tableau ondersteunt **de Athena-dataservice van Amazon** om verbinding te maken met Amazon S3, en beschikt over verschillende tools voor een rechtstreekse verbinding met NoSQL-databases. Voorbeelden van NoSQL-databases die vaak worden gebruikt met Tableau, zijn onder andere **MongoDB**, **Datastax** en **MarkLogic**.

**Databricks**, een platform voor datawetenschap en engineering, biedt dataverwerking op Spark, een populaire engine voor zowel batchgerichte als interactieve, uitbreidbare dataverwerking. Via een eigen connector met Spark kun je de resultaten van complexe machine learning-modellen van Databricks visualiseren in Tableau.



## Queryversnelling

Natuurlijk kun je machine learning toepassen en sentimentanalyses uitvoeren op big data, maar de eerste vraag die mensen vaak stellen, is toch: hoe snel is de interactieve SQL? SQL is uiteindelijk hét kanaal voor zakelijke gebruikers die big data willen gebruiken voor snellere, beter herhaalbare KPI-dashboards en verkennende analyses.

Deze behoefte aan snelheid heeft de invoer van snellere databases ingezet, waarbij in-memory en MPP (Massive Parallel Processing)-technologie wordt gebruikt, zoals **Exasol** en **MemSQL**, Hadoop-stores zoals Kudu en technologieën waarmee snellere query's met pre-processing mogelijk zijn, zoals **Vertica**. SQL-on-Hadoop-engines zoals Apache Impala, Hive LLAP, Presto, Phoenix en Drill, en OLAP-on-Hadoop-technologieën zoals **AtScale**, **Jethro Data** en Kyvos Insights zorgen ervoor dat query's sneller worden afgehandeld en dat het verschil tussen traditionele warehouses en big data nog kleiner wordt.

## Datacatalogus

Zakelijke datacatalogi vormen in feite een soort zakelijke woordenlijst van databronnen en algemene datadefinities, zodat het voor gebruikers gemakkelijker is om de juiste data te vinden om vanuit beheerde en goedgekeurde databronnen onderbouwde beslissingen te nemen. Ze zijn opgebouwd met metadata uit tabellen, weergaven en opgeslagen procedures door middel van het scannen van opgenomen databronnen. Bij het conserveren van data kunnen zelfs knowledge base-gegevens en weblinks worden meegenomen om gebruikers meer inzicht te geven in de context van de data, zodat er slimmer kan worden geclassificeerd en het ontdekken van data automatisch kan worden uitgevoerd.

Datacatalogi bestaan binnen visuele-analyse-oplossingen, maar zijn ook beschikbaar als standalone-optie voor een naadloze integratie met Tableau. Enkele van onze datacataloguspartners zijn **Informatica**, **Alation**, **Unifi**, Collibra en Waterline.

## Big data-analysearchitecturen

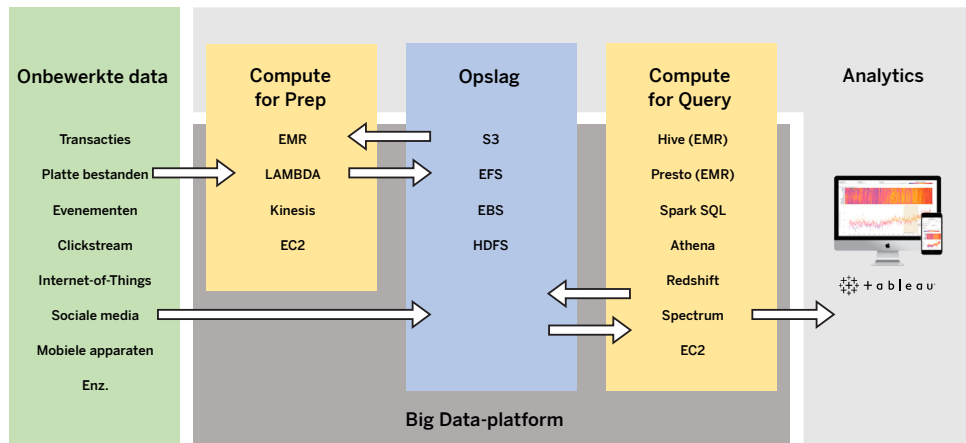
Het is belangrijk om te weten dat er geen uniforme aanpak is voor succesvolle big data-architecturen. Onze klanten hebben unieke, op maat gemaakte oplossingen voor hun big data-analyses, met verschillende platformen en tools voor de datapijplijnen. Desondanks is er wel sprake van een aantal overeenkomsten binnen deze architecturen, welke hebben bijgedragen tot het succes van deze big data-analyseplatformen.

**Disclaimer:** *Let op dat de volgende voorbeelden interpretaties zijn van Tableau en dat deze niet zijn ontworpen door de cloudproviders of klanten die zij vertegenwoordigen. Indien beschikbaar zijn de links naar de oorspronkelijke illustraties toegevoegd. Deze diagrammen zijn vereenvoudigde generalisaties met als doel het markeren van overeenkomsten in de belangrijkste elementen van de verschillende werkstromen. Ze zijn mogelijk geen weergave van elk onderdeel van het volledige big data-analyseplatform en representeren wellicht alleen bepaalde gebruikstoepassingen. Let tevens op dat 'compute for prep' het meest lijkt op 'process/catalog', en dat 'compute for query' het meest lijkt op 'analyze/model'.*

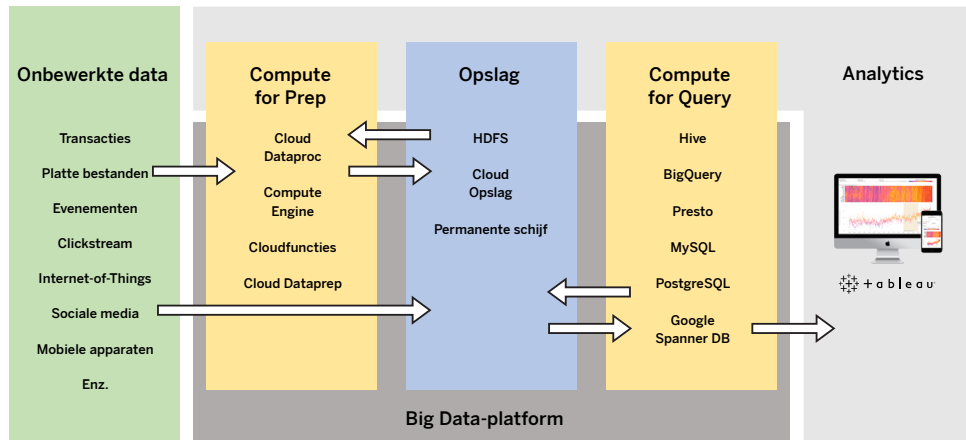


# Voorbeelden van grote cloudproviders

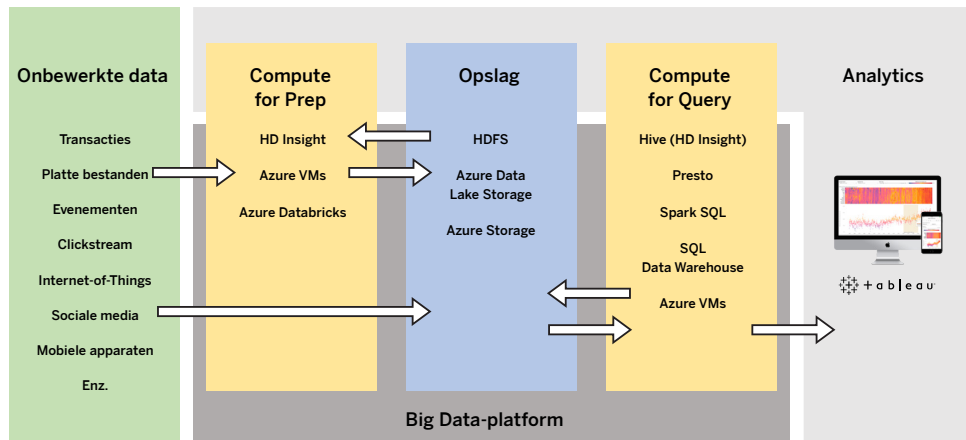
## Amazon Web Services



## Google Cloud Platform

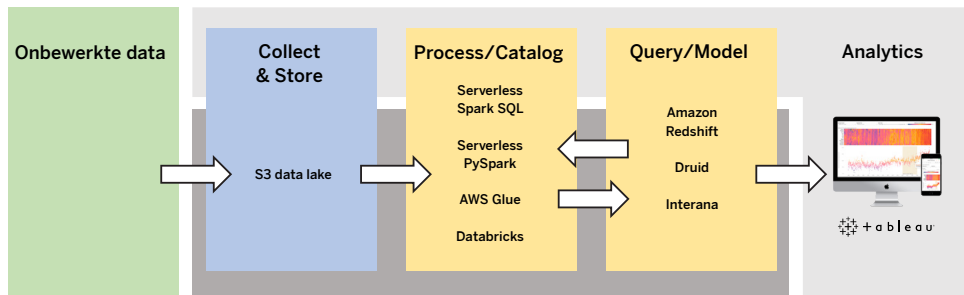


## Microsoft Azure

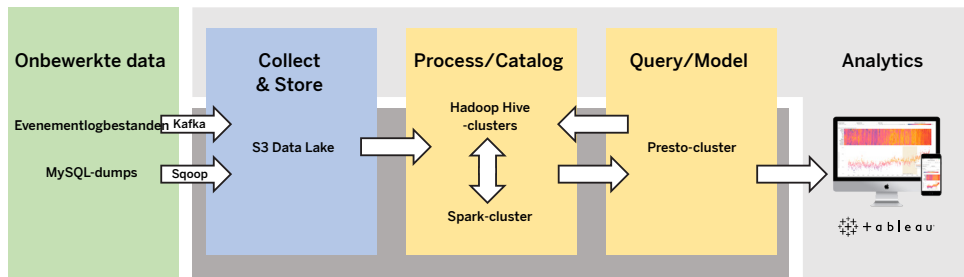


# Voorbeelden van Tableau-klanten

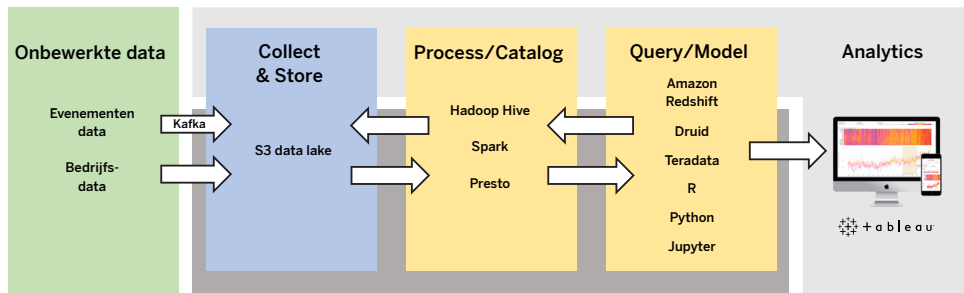
## Edmunds – [Meer info](#)



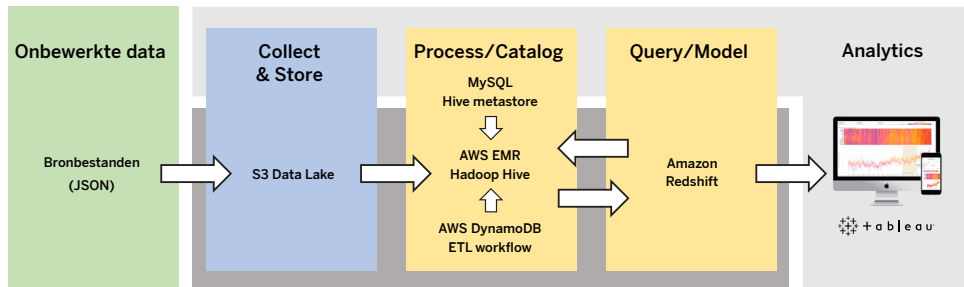
## Airbnb – [Meer info](#)



## Netflix – [Meer info](#)



## Expedia – [Meer info](#)



## Veelvoorkomende patronen

Ondanks dat geen twee bedrijfsarchitecturen hetzelfde zijn, kan overeenkomende patronen naast elkaar zetten je helpen bij het ontwikkelen van een strategie voor je eigen big data-analyseplatform. Dit is zijn de door ons opgemerkte overeenkomsten binnen succesvolle big data-analyse-architecturen:

- **Een opslaglaag** — Wordt meestal een data lake genoemd. Je datastrategie kan meerdere opslagomgevingen vereisen, maar sowieso moet gestructureerde, deels gestructureerde en ongestructureerde data aankunnen.
- **Compute-engines met en zonder server** — Een deel voor datavoorbereiding en analyse, en een deel voor het uitvoeren van query's. De dynamische aard van serverloos rekenen biedt meer flexibiliteit en elasticiteit, omdat er geen bronnen vooraf hoeven te worden toegewezen.
- **Ondersteuning voor volume, snelheid en variëteit** — Dit geldt niet alleen voor data, maar voor de groeiende complexiteit en het aantal gebruikerstoepassingen, waarvan een nog niet eens zijn ontdekt.
- **De juiste tool voor de taak** — Het is belangrijk om de componenten van je architectuur aan te passen aan je unieke datastrategie, maar het is ook belangrijk om flexibel te blijven in geval van een veranderende bedrijfsbehoefte.
- **Beheer en beveiliging op bedrijfsniveau** — Hoewel we niet dieper in zijn gegaan op deze gebieden, zijn beveiliging en beheer van fundamenteel belang voor een betere schaalbaarheid en een goed gebruik van je data.
- **Prijsbewust** — Houd rekening met de kosten bij het inschatten van de benodigde kracht en flexibiliteit van je big data-architectuur. De cloud biedt een heleboel ruimte om te groeien, maar het is belangrijk om rekening te houden met de financiële implicaties van je dataopslag en verwerking, simultaneïteit, latentie, gebruikstoepassingen voor analyses etc.

Het big data-landschap blijft ontwikkelen, maar er is absoluut sprake van een terugkerend thema:

Bedrijven moeten een algemeen, modern analyseplatform kunnen gebruiken om toegang te krijgen tot hun data, ongeacht of het groot of klein is en waar het zich bevindt. Met het juiste platform, de juiste processen en geschikte programma's voor mensen om mee te werken, zal datagestuurde besluitvorming van onschatbare waarde blijken.





## Over Tableau

Tableau is een compleet, gebruiksvriendelijk, visueel business intelligence-platform voor het bedrijfsleven, waarmee middels data supersnelle selfservice-analyses op inzicht kan worden verkregen in data. Op locatie of in de cloud, voor Windows of Linux: Tableau maakt gebruik van je bestaande technologische investeringen en groeit met je mee naarmate je dataomgeving verandert en groeit. Ontketen de kracht van je meest waardevolle activa, namelijk je data en je medewerkers.

## Aanvullende resources

[Bouwstenen voor een modern analyseplatform](#)

[Tableau bedrijfsanalyses: mogelijk gemaakt door IT](#)

[Tableau voor het bedrijf: IT-overzicht](#)

[Gratis proefversie van Tableau](#)