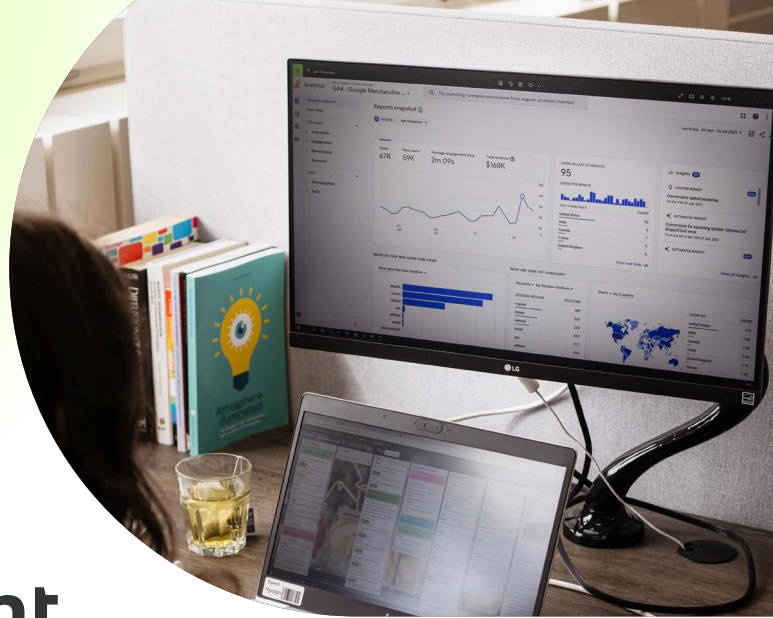


6 stappen van data naar inzicht met process mining



Door gebruik te maken van process mining kun je de daadwerkelijke paden die 'bewandeld' worden op jouw website visualiseren. Volg deze zes stappen om je (online) customer journeys inzichtelijk te maken.

1 Bepaal je doel

Ga eerst na welke data je nodig hebt om antwoord te geven op je onderzoeksvraag. Op je website zijn er verschillende typen verkeer aanwezig, zoals oriënterende gebruikers en gebruikers met een grote aankoopintentie. Bij het maken van je segmenten dien je hier rekening mee te houden.

Het kan interessant zijn om verschillende groepen (segmenten) met elkaar te vergelijken. Zo kun je inzichtelijk maken hoe een journey van een oriënterende gebruiker verschilt van de journey van iemand die overgaat tot een aankoop.

2 Bepaal je niveau

Hoe specifiek of generiek wil je een journey in kaart brengen? Je kunt een 'gemiddelde' flow van je website visualiseren om inzicht te krijgen in de meest voorkomende pijnpunten. Het is dan niet direct te herleiden wie deze pijnpunten ervaren.

3 Verzamel de data

Om gebruik te maken van process mining, heb je minimaal de data behorend tot de volgende variabelen nodig:

- ID (bijvoorbeeld visitID)
- Actie (bijvoorbeeld pagina)
- Timestamp

Daarnaast kun je zelf nog segmenten aanmaken. In deze segmenten kun je, afhankelijk van je doel, bepaalde gebruikers selecteren. Denk hierbij aan nieuwe gebruikers of gebruikers van een specifieke campagne.

De data om deze journeys inzichtelijk te maken d.m.v. process mining, kun je halen uit je analytics tool, bijvoorbeeld Google Analytics of Adobe Analytics.

Op de volgende pagina staan een aantal methoden om via Adobe Analytics en Google Analytics de data te krijgen die je nodig hebt.

Op basis van hoe vaak je dit doet en welke kennis je hebt, kun je hier een keuze uit maken.

Google Analytics (GA)

Ga eerst na welke data je nodig hebt om antwoord te geven op je onderzoeksvraag. Op je website zijn er verschillende typen verkeer aanwezig, zoals oriënterende gebruikers en gebruikers met een grote aankoopintentie. Bij het maken van je segmenten dien je hier rekening mee te houden.

Het kan interessant zijn om verschillende groepen (segmenten) met elkaar te vergelijken. Zo kun je inzichtelijk maken hoe een journey van een oriënterende gebruiker verschilt van de journey van iemand die overgaat tot een aankoop.

GA

Custom Report

Wanneer je gelimiteerde data hebt, kun je makkelijk via een custom report je data exporteren. Dit doe je door een custom report te maken met bijvoorbeeld de volgende variabelen; ID – Pagina – Timestamp - Paginaweergaves. Hierbij kies je ook het gewenste segment (zie stap 1).

Reporting API

Je kunt er ook voor kiezen om de data via de Reporting API op te halen. Dit is meer werk om op te zetten, maar zal zich later uitbetalen doordat je makkelijker meer data uit de tooling kunt halen en het ook voor andere zaken kan gebruiken.

Hier zijn ook plugins voor zoals Analytics Edge in Excel.

BigQuery

Voor GA 360 gebruikers is er nog de optie om de data te ontsluiten via BigQuery.

Via BigQuery heb je toegang tot bijvoorbeeld sessionId, visitorId en andere identifiers die je opslaat.

Adobe Analytics (AA)

Bij Adobe heb je ook niet standaard toegang tot de data die je nodig hebt op het juiste detailniveau. Hiervoor zal je ook het e.e.a. moeten meenemen in je implementatie.

Bij Adobe heb je standaard geen timestamps beschikbaar en zul je die altijd zelf moeten implementeren.

AA

Workspace

Op basis van gelimiteerde data exporteer je een tabel vanuit de Adobe Workspace om de analyse mee te doen.

Adobe API

Je kunt er ook voor kiezen om de data via de Adobe Analytics API V1.4 ontsluiten via bijvoorbeeld een package in R of Python.

Reportbuilder in Excel

Met Reportbuilder kun je al wat meer data exporteren. Echter zit je hier aan een maximum van 2 dimensies excl. granulariteit.

Data Warehouse

Via Data Warehouse heb je de mogelijkheid om zowel geaggregeerde data als niet geaggregeerde data op te halen. Hiermee kun je dus ook zelf het visitorID, visitNumber, en de benodigde eVar/prop identifiers inladen. Hier heb je wel een ftp server voor nodig.

Daarnaast moet je ook een timestamp hebben gedefinieerd in een prop (of eVar op hit niveau), want DataWarehouse heeft data standaard op uur niveau. Wil je grote hoeveelheden data ophalen, dan moet je dit middels een FTP export doen. Export richting mail heeft een maximum van 50 MB.

4 Tool of code?

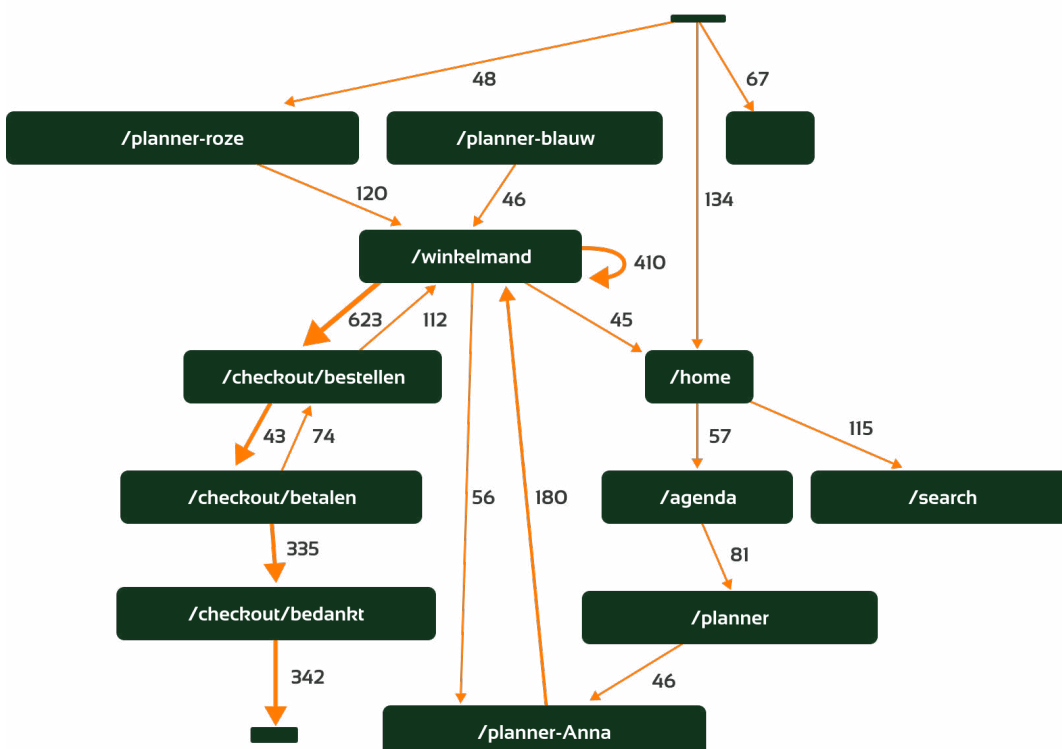
Wanneer je je data hebt verzameld en geprept, kun je beginnen met process mining. Er zijn diverse process mining tools die je kunt gebruiken.

Een voorbeeld van een process mining tool is ProM Lite. Dit is een process mining tool van de Technische Universiteit van Eindhoven. Voordeel van deze tool is dat deze gratis te gebruiken is, veel functionaliteit heeft en open source is. Nadeel is dat deze tool niet heel gebruiksvriendelijk is.

Andere opties zijn Disco en Celonis. Deze tools zijn gebruiksvriendelijker, maar wel prijzig. Tot slot kun je met packages in R/Python werken. De keuze is afhankelijk van je eigen situatie. Wij hebben gewerkt met ProM Lite.

5 Process mining in ProM Lite

De tool ProM Lite biedt vele functionaliteiten. In deze tool kun je een CSV bestand (download GA) omzetten tot een XES bestand. Vervolgens kun je het XES bestand gebruiken voor 'Heuristic Mining'. Heuristic Mining is een techniek die kan worden gebruikt om veel voorkomend gedrag te visualiseren. Vervolgens kun je kiezen voor een 'Directly-follows' graph. De output ziet er zo uit:



6 Profit - process mining insights

Zoals je ziet, zorgt de tool ervoor dat je direct een visueel overzicht krijgt van de meest bewandelde paden op de website. Daarnaast is er in de tool een optie waarbij je kunt aangeven hoe vaak een bepaald pad gewandeld moet zijn voor het wordt opgenomen in de visualisatie. In ProM Lite is dit de 'Frequency' balk. Hierdoor is het mogelijk om op een zeer gedetailleerd niveau te analyseren welke paden mensen nemen. Zo kun je goed zien welke onverwachte routes men neemt.

Door iets meer uit te zoomen krijg je een goed beeld bij welke paden vaker genomen worden. Hierdoor weet je meer over waar je het meeste impact kunt maken.

De visualisatie vanuit ProM kun je dus gebruiken om aan te tonen welke daadwerkelijke routes er worden gevolgd op jouw website. Een waardevol inzicht als je de online journey in kaart wilt brengen en/of wilt optimaliseren.



DIGITAL POWER

www.digital-power.com

Van data naar doen met Digital Power, jouw datapartner

Digital Power is een team van specialisten met passie voor data. Samen helpen we organisaties datagedreven te werken. Je vindt bij ons expertise op het gebied van Data Analytics, Technical Web Analytics, Customer Experience, Data Science en Data Engineering. Heb je een data vraagstuk, uitdaging of kans? We gaan graag voor – en mét – je aan de slag.

