

[Data analysis] 고급 google 애널리틱스 - 1.데이터 수집 및 처리 (정리) — 나무늘보의 개발 블로그

노트북: 첫 번째 노트북

만든 날짜: 2021-06-30 오후 11:35

URL: <https://continuous-development.tistory.com/249>

---

Data scientist/Data analysis

# [Data analysis] 고급 google 애널리틱스 - 1. 데이터 수집 및 처리 (정리)

2021. 6. 30. 18:06 수정 삭제 공개



## 1.데이터 수집 및 처리

차례

1. Google 애널리틱스 데이터 수집
2. 사용자 및 세션 분류하기
3. 구성 설정 적용하기
4. 데이터 저장하기 및 보고서 만들기
5. 측정 계획 수립하기

## 1. google 애널리틱스 데이터 수집

데이터를 수집하기 위해 자스크립트 추적 코드로 부터 시작

웹사이트에서 사용자 상호작용이 발생할때 마다 애널리틱스 추적 코드가 조회를 통해 google 애널리틱스로 전송 된다.

```
https://www.google-analytics.com/collect?v=1&_v=j43&a=1222585067&t=pageview&_s=1&dl=https%3A%2F%2Fwww.googleme  
rchandisestore.com%2F&dr=https%3A%2F%2Fwww.google.com%2F&ul=en-us&de=UTF-8&dt=Google%20Online%20Store&sd=24  
-bit&sr=1440x900&vp=1375x447&je=0&fl=21.0%20r0&_utma=41230772.810117230.1459896990.1464714869.1464802254.8&_u  
tmz=41230772.1464802254.8.7.utmcsr%3Dgoogle%7Cutmccn%3D(organic)%7Cutmcmd%3Dorganic%7Cutmctr%3D(not%2520p  
rovided)&_utmht=1464805887520&_u=SCCCgAAr~&jid=&cid=810117230.1459896990&tid=UA-30481-1&gtm=GTm-CGSQ&cd55=  
Retail&z=1860796108
```

이 안에 다양한 정보가 들어간다.

ex) 언어, 사용중인 페이지 이름, 화면 해상도 등등

### 조회의 3가지 유형

- 페이지 뷰 - 추적코드가 삽입된 웹페이지가 로드될 때 실행
- 이벤트 - 웹사이트의 특정 요소와 상호작용하는 매 순간을 추적(이벤트 액션, 카테고리, 라벨, 값 등의 데이터 매개변수가 전송)
- 거래 - 구매한 제품, 거래 ID 재고보관 단위 와 같은 전자상거래 구매 관련 데이터 전송

## 2.사용자 및 세션으로 데이터 분류하기

### 신규 사용자와 재사용자 구분

사용자 방문시 사용자의 브라우저 쿠키와 연결되는 임의의 고유 ID 생성  
이것을 통해 고유 사용자인지를 판별

### 제한사항

쿠키 삭제 또는 차단시 삭제 되어서 새로운 고유 ID 받게됨  
여러 세션에 걸쳐 사용자를 인식 할 수 있음

### 조회를 세션으로 분류

Google 애널리틱스에서는 웹사이트에서의 사용자 참여 수준을 파악하기  
위해 조회의 생성 시간에 기초해 사용자 조회를 분류합니다. 이러한 시간  
을 측정할 때 사용하는 측정항목은 '세션수'입니다.

ex)

- 홈페이지 방문 - > 세션 1 + 조회 1
- 홈페이지 방문 + 동영상 재생 -> 세션1 + 조회2
- 홈페이지 방문 + 30분 후 다시 영상재생 -> 세션 2 + 조회 2

기본 설정에서는 비활동시간이 30분 이상이면 세션 끊기고 이후에 행동  
할 시 새로운 세션을 발급 받는다.

추적코드의 데이터를 다른 데이터 소스와 결합

- 자동으로 조회를 전송하는 추적 코드

- Google 이외의 시스템으로부터 데이터를 수집하려면 URL 문자열을 통해 데이터 수집 조회를 직접 전송필요

### 3.구성 설정 적용하기

데이터 필터 , 목표, 데이터 분류, 맞춤 측정기준, 맞춤 측정항목, 데이터 가져오기 등의 기능을 구현하여 보고서에서 데이터를 더욱 세밀하게 정의하고 분석 할 수 있다.

- 데이터 필터 - 데이터 제외 , 특정 데이터 포함, 데이터를 수정
- 목표 - 도착, 이벤트 , 기간, 세션당 페이지
- 채널 그룹 및 콘텐츠 분류 - 데이터를 맞춤 채널로 분류할 수 있으며, 콘텐츠 분류를 이용하면 웹사이트의 구성에 따라 보고서에서 측정항목을 통합할 수 있다.
- 맞춤 측정기준 및 측정항목 - 측정기준과 측정항목을 직접 만들 수 도 있음
- 데이터 가져오기 - 웹사이트에서 수집한 조회 데이터와 오프라인 데이터를 결합

### 4.데이터 저장하기 및 보고서 만들기

데이터에 구성 설정이 적용되면 google 애널리틱스가 데이터를 측정기준으로 변환하고, 이와 연결된 측정항목을 계산하며, 신속한 검색이 가능하도록 자체 통합 데이터베이스 표에 각 측정항목을 저장한다.

## 측정항목

- 통합 집계 - 총 세션수, 사용자 수, 페이지 뷰 수 등
- 측정 기준에 따라 집계 - 세션수, 국가당 신규 사용자 수 등

## 주요 측정항목의 집계 방식

- 페이지뷰 조회의 시각을 포착한 후 이를 다음 페이지뷰 조회의 시각에서 빼서 '페이지에 머문 시간'을 계산합니다.
- '세션당 페이지수'는 사용자가 세션에서 발생시킨 순 페이지뷰 조회수의 평균값입니다.
- 평균 세션 시간은 사용자가 사이트를 떠나거나 세션의 제한 시간이 종료되기 전에 첫 번째 조회부터 마지막 조회까지 이어진 평균 시간입니다.
- 이탈률은 세션 시간이나 페이지에 머문 시간을 계산하기 위해 사이트에서 두 번째 상호작용 없이 상호작용을 한 번만 발생시킨 사용자를 분석하는 과정에서 계산됩니다. 이탈이 발생하면 이 방문의 페이지뷰에 할당된 세션 시간과 페이지에 머문 시간이 0이 됩니다.

## 측정기준과 측정항목 3개의 범위

- 조회 수준
- 세션 수준
- 사용자 수준

## 5.측정 계획 수립하기

수집할 데이터와 설정할 기능을 정하려면 비즈니스의 측정 전략을 명확하게 수립해야 함

- 거시 전환 - 비즈니스의 목표를 충족하는 핵심 행동(구매 등)
- 미시 전환 - 거시전환에 근접하도록 유도하는 행동(이메일 쿠폰 신청, 신제품 알림 신청)
- 전자 상거래 사이트 - 제품 구매
- 리드 생성 사이트 - 연락처 작성
- 콘텐츠 게시자 - 일정량의 콘텐츠 참여
- 온라인 정보 및 지원사이트 - 문제 해결을 위한 지원 안내 과정

미시 전환수는 거시전환으로 이어지는 사용자의 행동을 더욱 정확히 파악할 수 있게 해주는 측정 항목이다.

### 'Data scientist > Data analysis' 카테고리의 다른 글

[Data analysis] 고급 google 애널리틱스 - 1.데이터 수집 및 처리 (정리)

[Data analysis] google Analytics - 데이터 수집 및 처리

[Data analysis] google 애널리틱스 소개\_1(디지털 애널리틱스가 중요한 이유)

[Data analysis] 그럼에도 데이터 분석은 불완전하다 (데이터 분석의 힘 chapter....

[Data analysis] Google Analytics 란? (기본적인 설명)

[Data analysis] 데이터는 어떻게 전략이 되는가 (데이터 분석의 힘 chapter.6)



나아무늘보

혼자 끄적끄적하는 블로그 입니다.