



AGROCLUSTER
RIBATEJO

Business Strategy

Innovation
Branding
Solutions
Marketing
Analysis
Ideas
Strategy
Management

WEB ANALYTICS

ANALISE OS SEUS DADOS PARA
MELHORAR AS DECISÕES EMPRESARIAIS

João Mendes Moreira, jmoreira@fe.up.pt
Pedro Campos, pcampos@fep.up.pt



E-AGRO
INNOVATION



E-AGRO
MARKETS

Cofinanciado por:

COMPETE
2020

PORTUGAL
2020





AGROCLUSTER
RIBATEJO

Business Strategy

Innovation
Branding
Solutions
Marketing
Analysis
Ideas
Business
Management

Estrutura do workshop

1. Introdução (PC+JM)

2. Web analytics

2.1 Definição de web analytics (PC)

2.2 A necessidade do web analytics (PC)

2.3 Fundamentos web analytics para empresas (PC)

2.4 Google analytics (PC)

3. Metodologia CRISP-DM para dados web (JM)

4. Recolha de dados (PC)

4.1 web crawling

4.2 cookies

4.3 logs

5. Técnicas de modelação de dados

5.1 Tipos de aprendizagem (JM)

5.2 sistemas de recomendação (JM)

5.3 text mining (JM)

5.4 análise de redes sociais (JM)

6. search engines (PC)

6.1 search engine marketing (SEM)

6.2 search engine optimization (SEO)

7. Casos de estudo (JM)



E-AGRO
INNOVATION



E-AGRO
MARKETS

Cofinanciado por:



UNIAO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional

1. Introdução

*The Web is more a social creation than a technical one.
I designed it for a social effect—to help people work
together—and not as a technical toy.*

Tim Berners Lee, Weaving the Web (Harper)

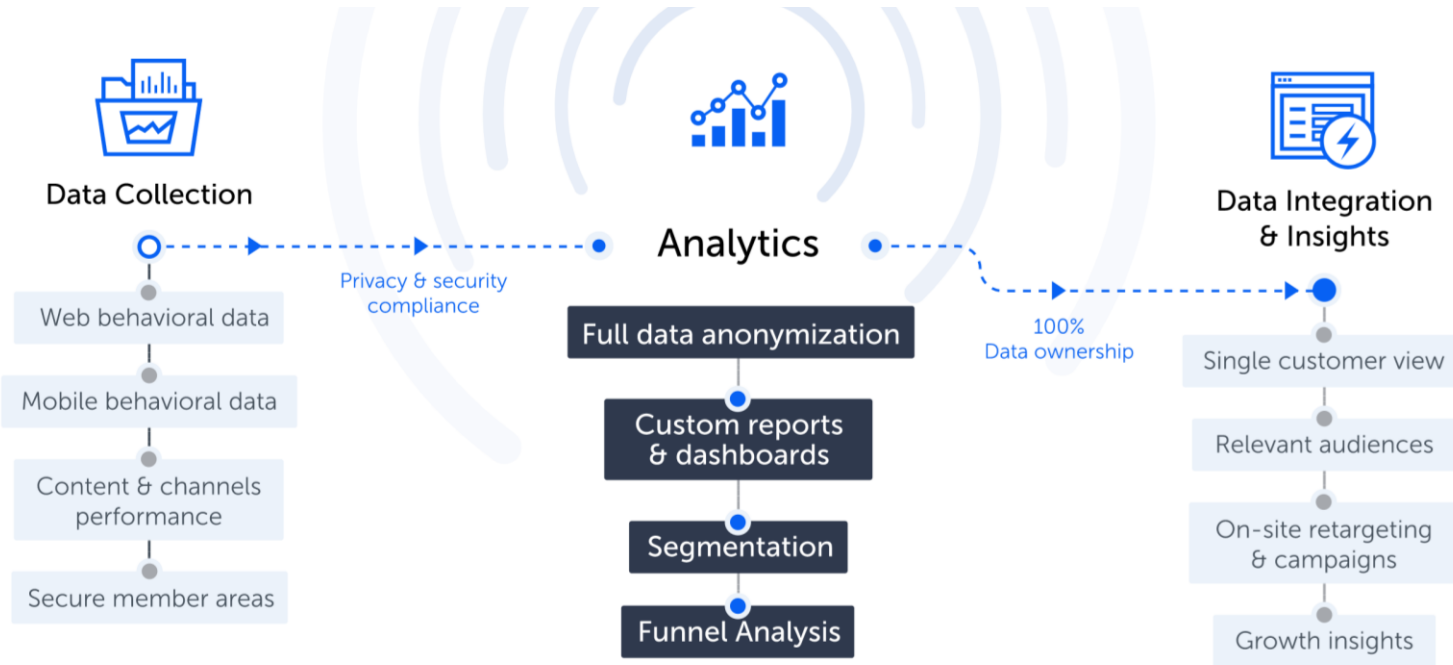
1. Introdução

- A Web é uma enorme fonte de informação.
- As informações encontram-se completamente desorganizadas.
- A quantidade de dados gerados e disponibilizados tem crescido exponencialmente e de forma contínua, dando origem ao paradigma do Big Data.
 - Existem 2,5 quintilhões de bytes de dados criados diariamente no nosso ritmo atual, mas esse ritmo está a acelerar com o crescimento da Internet das Coisas (IoT).
 - E admite-se que 90% dos dados existentes em todo o mundo foram gerados apenas nos últimos dois



Web analytics

1. Introdução



1. Introdução / Estrutura do Workshop

2. WebAnalytics, - principais métricas e fundamentos de web analytics e Google Analytics.

3. CRISP-DM: metodologia para na análise de dados que se adequa igualmente a projetos na área de Web Analytics.

4. Recolha de Dados: web crawling, cookies, logs

5. Modelação - Análise descritiva, Análise de diagnóstico, Análise preditiva e Análise prescritiva. São também introduzidos algoritmos de Sistemas de Recomendação, Análise de Texto e Análise de Redes Sociais.

6. Search Engine Marketing (SEM) e Search Engine Optimization (SEO).

2. Web Analytics



Definições

- O Web Analytics é a análise de dados qualitativos e quantitativos do web site e da concorrência, para promover uma melhoria contínua da experiência on-line dos utilizadores.

Usando algoritmos de machine learning, o software de Web Analytics encontra informações, sequências e padrões ocultos nos dados e utiliza-os para formar novas regras, agrupar utilizadores e prever o comportamento futuro dos clientes, transformando o “Big Data” (Delen, Sharda e Turban, 2015)

2. Web Analytics



Internet e WWW

- A Internet é o conjunto de ligações entre computadores operadas pelo governo indústria, academia, e empresas privadas.
- Apesar dos termos Internet e World Wide Web (WWW) serem usados de forma semelhante, a World Wide Web é apenas um dos muitos serviços que funcionam dentro da internet.
- Na verdade, a Web é um conjunto de documentos interconectados (as páginas web, ou que constituem os web sites) ligadas por hiperlinks e URLs.

2. Web Analytics



Métricas de Web Analytics

- As ferramentas gratuitas de análise de dados da Web - software que analisam o comportamento de visitantes do site - oferecem hoje em dia detalhes minuciosos sobre o quê, quando, onde, e por que razão os visitantes acedem aos web sites. Em particular, podem dar-nos informações sobre as métricas principais:
 - Acessos (hits)
 - Visualizações de página (pageviews)
 - Visitas (visits)
 - Visitantes únicos (unique visitors)
 - Referenciadores (referrers)
 - Palavras-chave e frases-chave (keywords and keyphrases).

2. Web Analytics



Métricas de Web Analytics

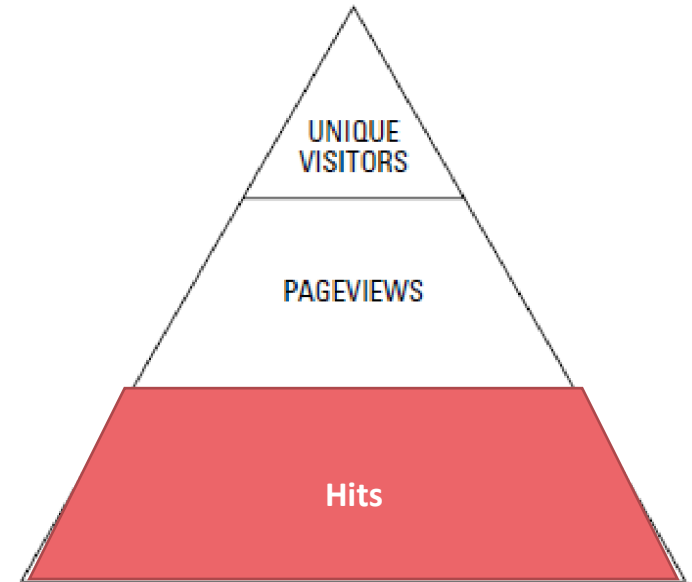
- Uma grande parte destes dados provêm do denominado clickstream ou sequência de cliques que segue a trajetória que um utilizador percorre ao clicar em qualquer objeto numa página web. Os dados de clickstream são registados em aplicações através de cookies.

Uma **métrica** é uma medida quantitativa que descreve eventos ou tendências num web site. Para ajudar a definir métricas, usam-se medidas de desempenho chave denominadas (KPI – Key Performance Indicators) que ajudam a avaliar o desempenho relativamente aos objetivos propostos pela empresa.

2. Web Analytics

Métricas de Web Analytics

- **Acessos (hits)** - esta é uma das métricas mais utilizadas em web analytics e uma das mais falaciosas, se for mal interpretada. Os acessos aos web sites são medidos através do número de ocorrências ou visualizações. As ocorrências ou visualizações são criadas quando o servidor Web entrega um ficheiro ao navegador (browser) de um visitante. Os ficheiros PDF, de som, Word, outros documentos e imagens são alguns exemplos de ficheiros que geram hits. Por exemplo, o pedido de uma página com cinco imagens contaria como seis hits: um hit para o página em si mais um hit para cada uma das cinco imagens. Esta tem sido um métrica popular em web sites que esperam pontuar publicidade, mas que, tal como referido, pode enganar a interpretação

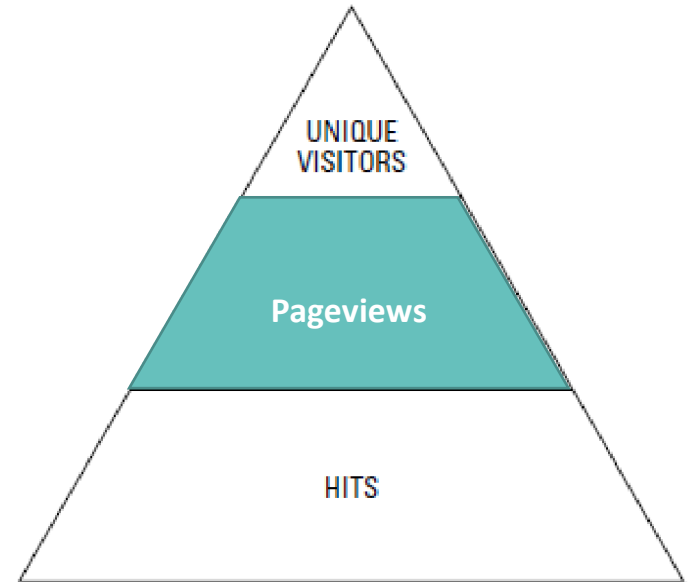


A pirâmide da terminologia de Web Analytics
Fonte: Sostre e LeClaire, (2007)

2. Web Analytics

Métricas de Web Analytics

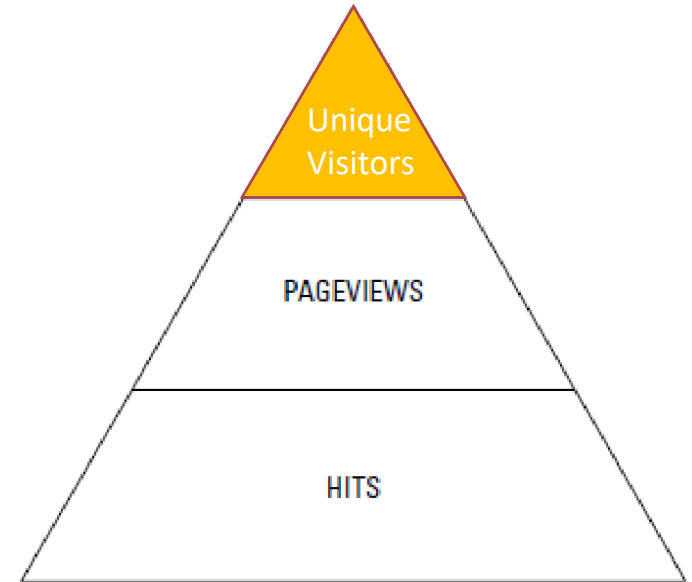
- **Visualizações de página (pageviews)**- Uma exibição de página é registada sempre que um utilizador visualiza uma página no web site. Os programas de web analytics dividem o número de visitantes pelo número de visualizações de cada página para determinar o número médio de páginas que cada visitante visualizou. Se esse número for baixo, isso é sinal de que talvez seja necessário repensar o conteúdo, o design ou a estrutura do web site.



2. Web Analytics

Métricas de Web Analytics

- **Visitas** - Às vezes também chamada de sessão ou sessão de utilizador, uma visita descreve a atividade de um utilizador individual no web site. Também se pode dizer que uma visita consiste numa série de visualizações efectuadas pelo mesmo visitante. É importante notar que a maioria das ferramentas de web analytics termina a sessão se o visitante permanecer inativo por 30 ou mais minutos, embora esse limite de tempo.
- **Utilizadores/Visitantes únicos** - A métrica de visitantes únicos representa o número de pessoas individuais que visita o web site. Cada indivíduo é contado apenas uma vez, pelo que se uma pessoa visitar um site diversas vezes no período de análise, esse comportamento pode contar como diversas sessões, embora apenas como um visitante único.



2. Web Analytics

Outras métricas, mais complexas,

- fornecem resultados importantes para a tomada de decisão

Taxa de conversão a conversão significa venda ou encomenda de um determinado bem ou serviço. A Taxa de conversão, expressa em termos de resultados a dividir pelo número de utilizadores únicos, mede a capacidade de venda de um web site. Esta métrica é especialmente interessante no comércio eletrónico.

Taxa de saída ou taxa de abandono (bounce rate) é o número de pessoas que sai do web site antes de executar uma tarefa importante, tal como uma conversão.

Lealdade, não existe um consenso claro quanto à forma de a medir, sendo que há quem use a duração da visita ao web site, a quantidade de importação de documentos, etc.

2. Web Analytics

Outras métricas, mais complexas,

Os valores de referência para estas métricas são os seguintes:

- **Taxa de rejeição** (bounce rate - idealmente abaixo de 60%)
- **Tempo no site** (acima de 60 segundos indica um visitante comprometido)
- **Visualizações de página** (quanto mais, melhor)
- **Páginas por visita** (quanto mais, melhor)

2. Web Analytics

Importância para as empresas

Os dados para web analytics podem ser gerados por diversas fontes:

- **Tráfego direto**: corresponde aos utilizadores do web site que entram porque digitam diretamente o endereço, ou usam uma URL que dá acesso direto ao web site.
- **Referenciador do site**: outros Web sites que enviam tráfego para o nosso site como resultados de campanhas, blogs, etc.
- **Motores de busca**: Google, Yahoo, Bing e outros, incluindo tráfego orgânico e pago.

2. Web Analytics

Importância para as empresas

Os dados para web analytics podem ser gerados por diversas fontes (cont.):

- **Outros:** Campanhas, e-mail, marketing direto, etc.
- **Social media**, por exemplo (Twitter, Facebook, LinkedIn, Instagram, etc.), são importantes fontes de informação que podem ser usadas para tomada de decisão numa empresa.

2. Web Analytics

Importância para as empresas

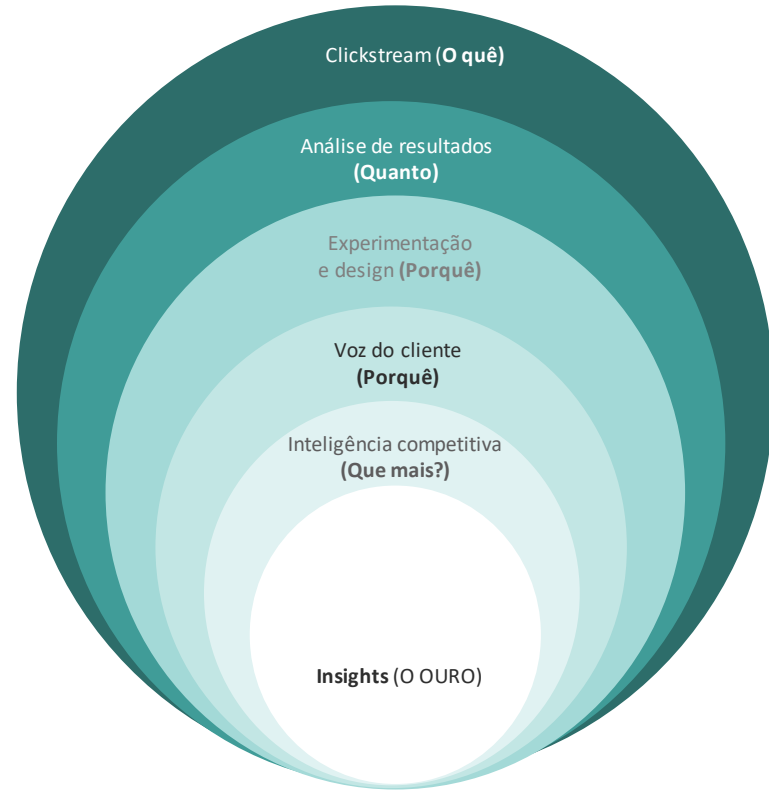
Muitas empresas dispõem de CRM que lhes permite organizar a informação de interação com os clientes proveniente de diversos canais: presencial, online (web), telefone, etc. Os CRM permitem ajudar a recolher dados úteis para web analytics.

O CRM é um tipo de sistemas de informação que se concentra na recolha, processamento e coordenação de informações sobre interações comerciais com clientes, incluindo contactos, transações comerciais, marketing, vendas e serviços. O uso do site de uma empresa pode fornecer diversas informações de marketing que po

2. Web Analytics

Importância para as empresas

Como gerir todos estes dados?



Fonte: adaptado de Alghalith (2015)

2. Web Analytics

Ferramentas de Web Analytics: o Google Analytics

- A maioria das empresas usa ferramentas do Google Analytics.
- Mas existem outras ferramentas, tais como Omniture Site Catalyst, Webtrends, Clicktracks ou Xiti para entender o que está acontecendo nos web sites.



2. Web Analytics



Ferramentas de Web Analytics: o Google Analytics

O Google Analytics é uma das ferramentas mais usadas de web analytics. Pertence ao universo Google e fornece estatísticas detalhadas do tráfego da web, sendo usada por mais de 60% dos proprietários de sites. Em termos técnicos, o Google Analytics contém linguagem javascript em cada página do site que o visitante aceder. Este procedimento permite enviar os dados ao Analytics, que, por sua vez, os transmite ao proprietário do web site.

2. Web Analytics



Ferramentas de Web Analytics: o Google Analytics

O Google Analytics permite:

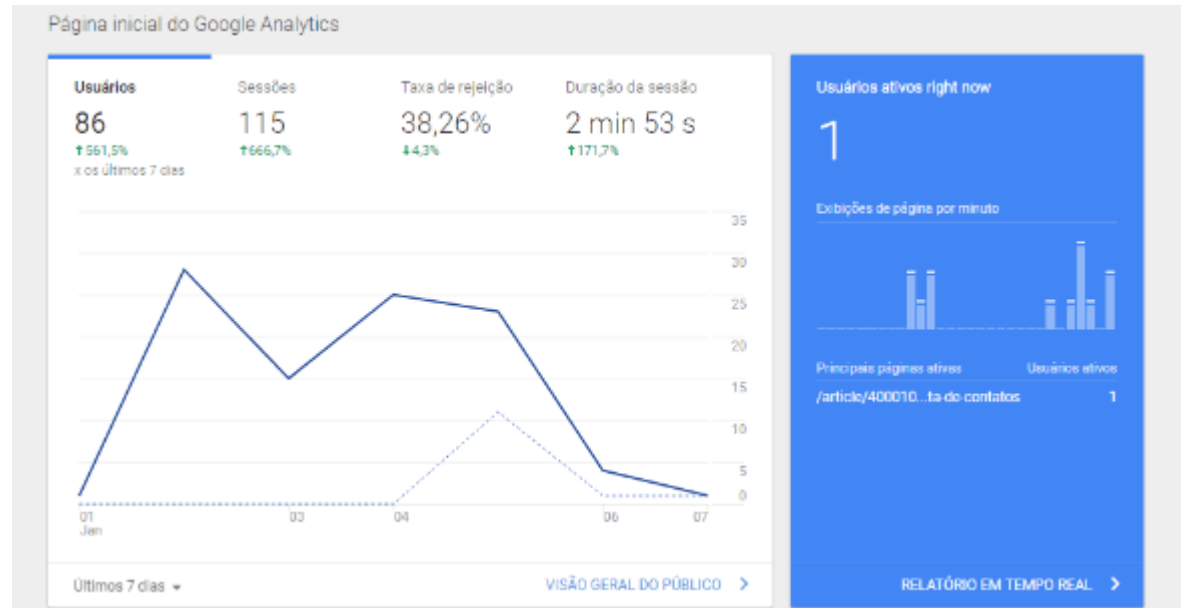
- analisar quais os dispositivos que as pessoas utilizam para navegar pelo site (telemóvel, computador desktop, tablet, etc.)
- qual o tempo médio de cada sessão
- calcular quais as páginas que geram mais leads para o negócio;
- descobrir quais as páginas do web site com mais ou menos visitas;
- determinar o que os utilizadores pesquisam no web site.
- ... e o que pesquisaram no Google para encontrar o web site.

2. Web Analytics



Ferramentas de Web Analytics: o Google Analytics

- O Google Analytics funciona como um painel de controlo (ou dashboard) para gerir as visitas a um web site. Através das informações que proporciona, é possível analisar, por exemplo, o desempenho de vendas de cada produto numa loja de e-commerce.



2. Web Analytics



Ferramentas de Web Analytics: o Google Analytics

- Mais do que isso, é possível saber que tipo de ação de marketing digital está a gerar os melhores resultados, assim como o valor desse retorno. O Google analytics tem templates pré-definidos de web analytics contendo as seguintes métricas principais para medir o desempenho de um web site, tais como taxa de rejeição (bounce rate), duração da sessão, Visualizações de página e Páginas por visita.

🔍 Pesquisar relatórios e ajuda

Descrição geral do público-alvo ✓

📄 GUARDAR 📄 EXPORTAR ↵ PARTILHAR 📊 ESTATÍSTICAS

- 🏠 Página inicial
- 🔧 Personalização
- RELATÓRIOS
- 🕒 Tempo real
- 👤 Público-alvo

Descrição geral

Utilizadores ativos

Valor do cliente
BETA

Análise da coorte
BETA

Públicos-alvo

Explorador de utilizadores
 Dados demográficos

🔗 Atribuição BETA

💡 Descobrir

⚙️ Administração

🕒 Todos os utilizadores
 100,00% Utilizadores

🕒 + Adicionar segmento

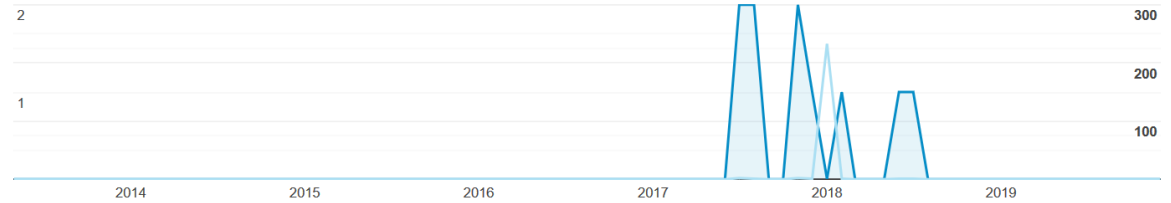
01/05/2013 - 23/12/2019 ▾

Descrição geral

Páginas/Sessão vs. Sessões

De hora a hora Dia Semana Mês

● Páginas/Sessão ● Sessões



Utilizadores
239



Novos Utilizadores
239



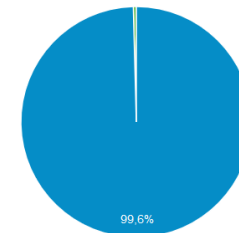
Sessões
242



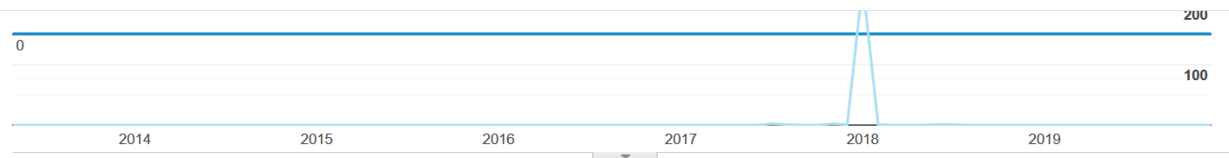
Número de sessões por utilizador
1,01



■ New Visitor ■ Returning Visitor



- 🔍 Pesquisar relatórios e ajuda
- demográficos
 - Interesses
 - Geográficos**
 - Comportamento
 - Tecnologia
 - Dispositivos móveis
 - Descrição geral**
 - Dispositivos
 - Entre dispositivos BETA
 - Personalizado
 - Testes de referência
 - Fluxo de utilizadores
 - Aquisição
 - Comportamento
 - Conversões
 - Atribuição BETA



Dimensão Principal: Categoria do Dispositivo

Inserir linhas Dimensão secundária Tipo de Ordenação: Predefinição 🔍 avançado 📄 ⏪ ⏩ 🗑️

Categoria do Dispositivo ?	Aquisição			Comportamento			Conversões	
	Utilizadores ?	Novos Utilizadores ?	Sessões ?	Taxa de rejeições ?	Páginas/Sessão ?	Duração média da sessão ?	Taxa de conversão de objetivos ?	Objetivos Concluídos ?
	0 <small>% do total: 0,00% (0)</small>	238 <small>% do total: 99,58% (239)</small>	239 <small>% do total: 98,76% (242)</small>	2,09% <small>Média por visualização de propriedade: 2,07% (1,26%)</small>	0,04 <small>Média por visualização de propriedade: 0,06 (-39,25%)</small>	00:00:04 <small>Média por visualização de propriedade: 00:00:10 (-59,64%)</small>	0,00% <small>Média por visualização de propriedade: 0,00% (0,00%)</small>	0 <small>% do total: 0,00% (0)</small>
<input type="checkbox"/> 1. desktop	0 (0,00%)	237 (99,58%)	237 (99,16%)	2,11%	0,02	00:00:00	0,00%	0 (0,00%)
<input type="checkbox"/> 2. mobile	0 (0,00%)	1 (0,42%)	2 (0,84%)	0,00%	2,00	00:07:50	0,00%	0 (0,00%)

Mostrar linhas: 10 Ir para: 1 1 - 2 de 2 ⏪ ⏩

Este relatório foi gerado em 24/12/19 às 15:42:16 - [Atualizar Relatório](#)

🔍 Pesquisar relatórios e ajuda

🏠 Página inicial

▸ 🗑️ Personalização

RELATÓRIOS

▸ 🕒 Tempo real

▾ 👤 Público-alvo

 Descrição geral

 Utilizadores ativos

 Valor do cliente BETA

 Análise da coorte BETA

 Públicos-alvo

 Explorador de utilizadores

 ▸ Dados demográficos

 ▸ Interesses

 ▸ Geográficos

🔗 Atribuição BETA



Dados demográficos

Idioma

País

Cidade ▸

Sistema

Navegador

Sistema operativo

Fornecedor de serviços

Google Mobile

Sistema operativo

Fornecedor de serviços

Resolução de ecrã

Cidade	Utilizadores	% Utilizadores
1. Atlanta	126	52,72%
2. New York	44	18,41%
3. (not set)	35	14,64%
4. San Francisco	27	11,30%
5. Samara	2	0,84%
6. Tippecanoe	1	0,42%
7. Riverside	1	0,42%
8. Horse Shoe	1	0,42%
9. Clayton	1	0,42%
10. Short Pump	1	0,42%

[ver relatório completo](#)

Este relatório foi gerado em 24/12/19 às 15:52:01 - [Atualizar Relatório](#)

2. Web Analytics



Anúncios online e o Google Analytics

- As ferramentas de web analytics permitem acompanhar diversas formas de anúncios online tais como banners, pop-ups, campanhas de email, etc.
- Para além destas formas de publicidade, o Google analytics permite também acompanhar as palavras chave de pesquisa num motor de busca: as keywords.
- A Google oferece uma ferramenta interessante - criador de URL - especificamente para os seus utilizadores do AdWords. Para se ter acesso, apenas é necessário preencher um formulário simples.

Se a conta do Google Analytics estiver associada a uma conta ativa do Google AdWords, será possível fazer análise das palavras pesquisadas, por quem, como e quando.

Estas abordagens fazem parte do Search Engine Marketing (SEM), uma prática de marketing que as empresas usam para promover os seus produtos/serviços na webpage de resultados de um motor de pesquisa.

2. Web Analytics



Anúncios online e o Google Analytics



- A publicidade paga nos motores de busca online constitui a área do mercado da publicidade online com maior dinâmica. Esse espaço disponível para os anúncios pagos com base em palavras chave (keywords), ou posições, é vendido através de leilões e o seu pagamento é calculado tendo em conta o número de clicks que cada posição recebe. A Google promove espaços para a inserção de campanhas publicitárias.
- É através do Google Ads (antigamente denominado Google AdWords) que essas campanhas são geridas. O Search Engine Marketing (SEM), também chamado de keyword advertising, foca-se em alocar os anúncios que aparecem nos resultados de pesquisa dos motores de busca às necessidades dos utilizadores.

2. Web Analytics



Web Analytics e a web social

- Existem outras diferentes de fazer web analytics que não se baseiam nas métricas definidas anteriormente (visitas, páginas visitadas, etc.).
- Trata-se de dados da web provenientes de social media (tais como Twitter, Facebook, Instagram, LinkedIn, etc.). Nestes casos, utilizam-se frequentemente aplicações do tipo API (Application Programming Interface) para obter os dados a partir dos quais e podem fazer vários tipos de análise..



2. Web Analytics



Web Analytics e a web social

- O Twitter, por exemplo, pode ser visto como um serviço de microblog que permite que as pessoas se comuniquem com mensagens curtas que correspondem aproximadamente a pensamentos ou ideias.
- Historicamente, esses tweets limitavam-se a 140 caracteres, embora esse limite tenha sido expandido e possa ser mudar novamente no futuro. Nesse sentido, podemos pensar no Twitter como algo semelhante a um serviço global de mensagens de texto de alta velocidade.

2. Web Analytics



Web Analytics e a web social

- A API to Twitter permite-nos obter informação sobre a rede de cada utilizador, usar medidas próprias para avaliar a sua posição na rede, etc.

The screenshot shows a web browser window with the address bar displaying 'developer.twitter.com/#modal-analyze'. The page content includes a purple icon with a checkmark, the heading 'Analyze', and the main text: 'Evaluate Twitter data to inform business decisions.' Below this, there is a paragraph of text: 'Twitter data is the most comprehensive source of live, public conversation worldwide. Our REST, streaming, and Enterprise APIs enable programmatic analysis of Tweets back to the first Tweet in 2006. Whether you are building a solution for brands, one for your own team, or conducting research—our endpoints enable rich insight into audiences, market movements, emerging trends, key topics, breaking news, and much more.' A link 'Learn more >' is visible at the bottom of the content area. The browser's taskbar at the bottom shows various application icons and the system clock indicating 03:02 on 14/01/2020.

2. Web Analytics



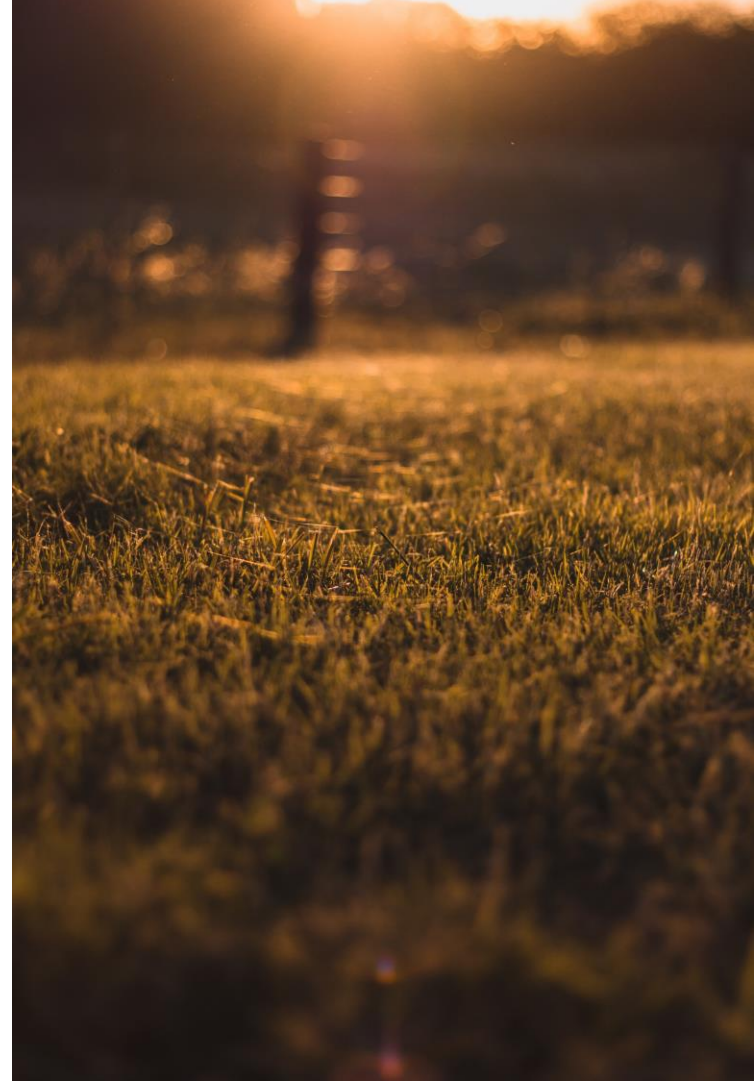
Web Analytics e a web social

- O mesmo se passa com o Facebook ou com o Instagram. Ou seja, existem diversas métricas que podem ser analisadas convenientemente neste tipo de dados.
- Russel (2019) apresenta diversas API para lidar com estas fontes de social media e usa Python e algoritmos de data mining para as analisar, tais como clustering ou análise preditiva

Description	Segmentation Available	
engagements	Total number of engagements	✓
impressions	Total number of impressions	✓
retweets	Total number of retweets	✓
replies	Total number of replies	✓
likes	Total number of likes	✓
follows	Total number of follows	✓
card_engagements	Total number of card engagements	

4. Recolha de Dados

- A **recolha de dados** é uma etapa fundamental em *Web Analytics*.
- A forma como os dados são recolhidos tem um **grande impacto** na qualidade dos dados finais.
- Existem várias formas de recolher dados da Internet, desde a extração de textos nas páginas web (**web crawling**), à possibilidade de instalação de mecanismos para obter informações dos utilizadores (**cookies**), até à recolha de dados para efeitos de caracterização dos utilizadores de um web site (**logs** e estatísticas dos motores de busca).



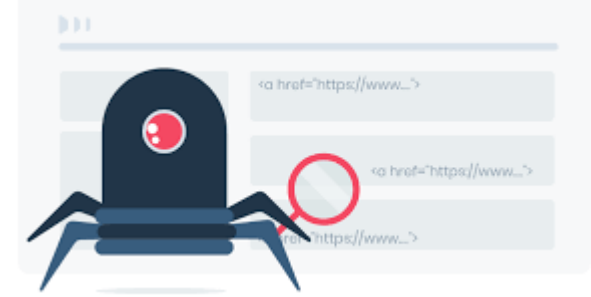
4. Recolha de Dados



Web Crawling

- O web crawler ou bot é um algoritmo utilizado para analisar o código de um website e recolher informações.
- Estas informações permitem classificar os dados e gerar novo conhecimento.
- Estes bots são usados pelos motores de busca como o Google, Bing ou outros de modo a apresentar a informação mais relevante aos utilizadores.
- Estes algoritmos percorrem toda a web recolhendo informações de todas as páginas dos sites de forma a apresentarem os resultados mais relevantes cada vez que alguém pesquisa num motor de busca.

4. Recolha de Dados



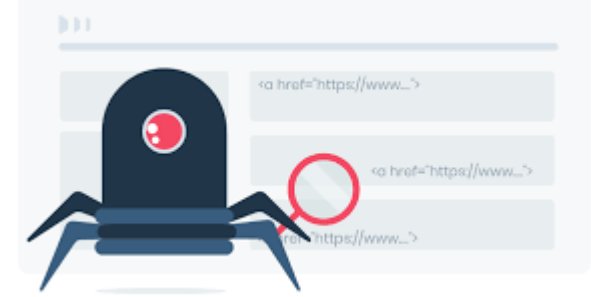
Web Crawling

Porque é importante o rastreamento web nos negócios? Através deste rastreamento pelos *bots* é possível saber o que é dito sobre a empresa nas notícias, em fóruns, blogs, pelos clientes, nas redes sociais, ou outros sites.

Com um rastreador web saberá se o conteúdo que produz tem bom desempenho ou se é necessário otimizar a nível de SEO para atrair e converter melhor o seu público.es



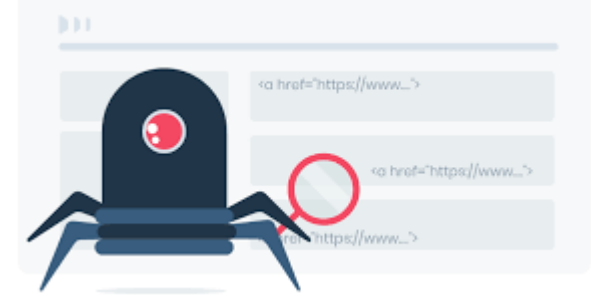
4. Recolha de Dados



Web Crawling

- O Google é um dos principais motores de busca e fontes de aquisição de visitas para a maioria dos sites. Por isso, é importante conseguir tirar o melhor proveito desse tráfego e gerar vendas.
- Através dos web crawlers é possível perceber quais as otimizações que podem ser feitas a nível de SEO para melhorar a experiência do utilizador e o site ser encontrado mais facilmente nos motores de busca.
- São mais de 200 os fatores que influenciam o ranqueamento do Google, desde tags, estrutura do código, navegação, velocidade de carregamento, dispositivos móveis, e muito mais, que poderá analisar com um software de rastreamento para que aumente o tráfego e conversões no seu site

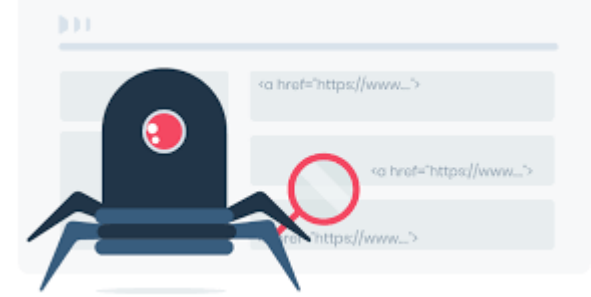
4. Recolha de Dados



Web Crawling

- Existem também ferramentas que ajudam os gestores a analisar o site da empresa de forma a obterem melhorias, alguns exemplos:
 - Oncrawl - bot que realiza auditorias de SEO no site.
 - Dyno Mapper - permite criar automaticamente mapas do site.
 - Screaming frog - que lhe permite melhorar o SEO no site.
 - Apify, permite-lhe analisar a concorrência e tomar decisões importantes no site.

4. Recolha de Dados



Web Crawling versus Web Scraping

Movimento:

- Web Scraping literalmente raspa todos os dados **selecionados** e retira-os.
- Web Crawling rastreia as informações de dados do destino selecionado.

Trabalho:

- Web Scraping é feita geralmente de forma manual.
- Web Crawling é feita com a ajuda de um agente de bot.

<https://www.aweb scraping.com/web-crawling-vs-web-scraping/>

4. Recolha de Dados

Cookies

- Quando visita um site através do computador ou dispositivos móveis é instalado um pequeno arquivo de texto (cookie) que permite durante algum tempo memorizar as ações ou preferências que tem no site, para que quando o volte a visitar não tenha que as definir novamente.
- Os cookies armazenam dados como idioma escolhido, nome do utilizador, tamanho dos caracteres, opções de guardar a palavra-passe, entre outros.
- Os cookies são específicos de cada navegador, ou seja, se armazenar cookies no Firefox estes não serão mantidos no Google Chrome, por exemplo.

Por norma, os cookies são utilizados para melhorar a experiência do utilizador no site e evitar a repetição de informação.



4. Recolha de Dados

Cookies

Quais os tipos de cookies que existem?

Existem diversos tipos de cookies de acordo com a finalidade a que se destinam, por exemplo:

- **Cookies sociais:** são utilizados para permitir que o utilizador partilhe o conteúdo dos sites nos seus perfis nas redes sociais.
- **Cookies publicitários:** permitem direccionar os anúncios do site em função dos interesses de cada utilizador, limitando a quantidade de vezes que este vê um anúncio no site.
- **Cookies analíticos:** são utilizados anonimamente para analisar estatísticas e melhorar o conteúdo do site.



4. Recolha de Dados

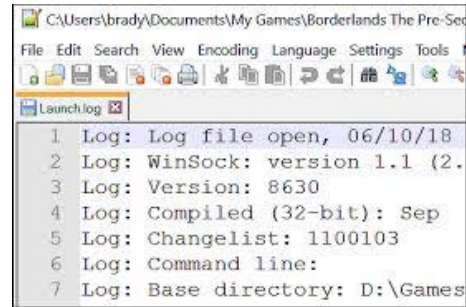
Cookies



- Os cookies não podem armazenar informações complexas, pois estão limitados a informações simples em cadeia. Existem muitas limitações no tamanho do texto do cookie. O cookie individual pode conter uma quantidade muito limitada de informações (não mais que 4 kb). O utilizador tem a opção de desativar os cookies na configuração do navegador em resposta às preocupações de segurança ou privacidade. No entanto, poderá existir problemas para sites que os exigem e os cookies não funcionam. Com a ativação de cookies os navegadores mantêm o controlo de todos os sites que visitou. Também sites terceiros podem aceder às informações armazenadas por esses cookies. Esses sites podem ser anunciantes, ou até outros utilizadores.

4. Recolha de Dados

Ficheiros de logs



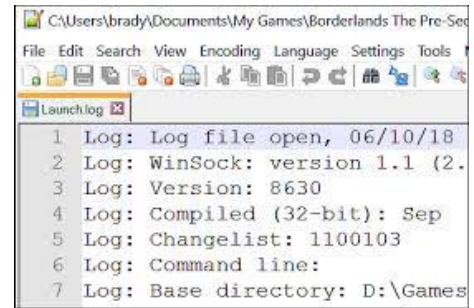
```
1 Log: Log file open, 06/10/18
2 Log: WinSock: version 1.1 (2.
3 Log: Version: 8630
4 Log: Compiled (32-bit): Sep
5 Log: Changelist: 1100103
6 Log: Command line:
7 Log: Base directory: D:\Games
```

- Na computação, um ficheiro de log regista eventos que ocorrem num software que está a ser executado.
- Para pesquisa na Web, um log de transações é um registo eletrónico de interações que ocorreram durante um episódio de pesquisa entre um mecanismo de pesquisa na Web (browser) e utilizadores que procuram informações nesse mecanismo de pesquisa na Web.

Os ficheiros de log contêm dados recolhidos pelo servidor de web. Trata-se de ficheiros que registam a atividade num software. Um servidor web regista a sua atividade num ficheiro de texto. Os analistas conseguem usar essa informação e produzir relatórios,

4. Recolha de Dados

Ficheiros de *logs*

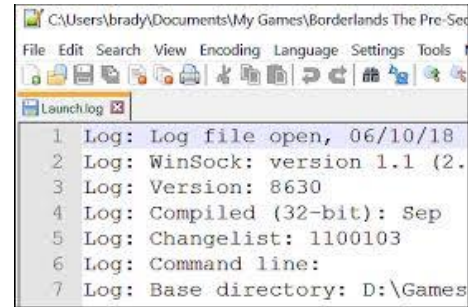


```
C:\Users\brady\Documents\My Games\Borderlands The Pre-Sec
File Edit Search View Encoding Language Settings Tools
Launch.log
1 Log: Log file open, 06/10/18
2 Log: WinSock: version 1.1 (2.
3 Log: Version: 8630
4 Log: Compiled (32-bit): Sep
5 Log: Changelist: 1100103
6 Log: Command line:
7 Log: Base directory: D:\Games
```

- Um log do servidor da web é um diário registado num servidor web que contém '**hits**' ou registos de todas as solicitações que o servidor recebeu.
- Os dados recebidos são armazenados anonimamente e incluem detalhes como a hora e a data em que a solicitação foi feita, o endereço IP da solicitação, o URL / conteúdo solicitado e o user agent do navegador.
- Esses ficheiros existem normalmente para auditoria e solução de problemas técnicos do site, mas também podem ser extremamente valiosos para a auditoria de SEO (Search Engine Optimization).

4. Recolha de Dados

Ficheiros de logs



```
1 Log: Log file open, 06/10/18
2 Log: WinSock: version 1.1 (2.
3 Log: Version: 8630
4 Log: Compiled (32-bit): Sep
5 Log: Changelist: 1100103
6 Log: Command line:
7 Log: Base directory: D:\Games
```

- A aparência de um ficheiro de logs depende do tipo de servidor, mas há elementos comuns entre todos os ficheiros:
 - Server IP
 - User-agent
 - Timestamp (date & time)
 - HTTP status code
 - Method (GET / POST)
 - Requested URL (aka: URL stem + URL query)
 - Referrer (the external site, such as Google, from which a user arrives)

Business Strategy

- Innovation
- Branding
- Solutions
- Marketing
- Analysis
- Ideas
- Business Management



AGROCLUSTER
RIBATEJO



E-AGRO
INNOVATION



E-AGRO
MARKETS

Cofinanciado por:



UNIAO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional