



Fundamentos da Coleta de Dados Online

Aprendendo os fundamentos da coleta de dados online



Índice

Introdução	3
A internet mudou a pesquisa de mercado	4
Fontes de amostras online	6
Métodos de amostragem	10
Amostragem por cotas online efetivas	14
Planejando a coleta de dados online	17
Pesquisas eficazes para o trabalho de campo online	19
Dados de qualidade	21
Privacidade	24
Relatórios e visualização	26



Introducción

Este ebook fornece orientação sobre como implementar a coleta eficaz de dados online destinada à pesquisa de mercado.

A coleta de dados começa com a definição do objetivo e do projeto de pesquisa, e termina quando os dados são disponibilizados para serem analisados. Neste contexto “eficaz” significa o justo equilíbrio entre custo, tempo e qualidade.

Este ebook pode te interessar se você se identifica com qualquer um dos seguintes perfis:

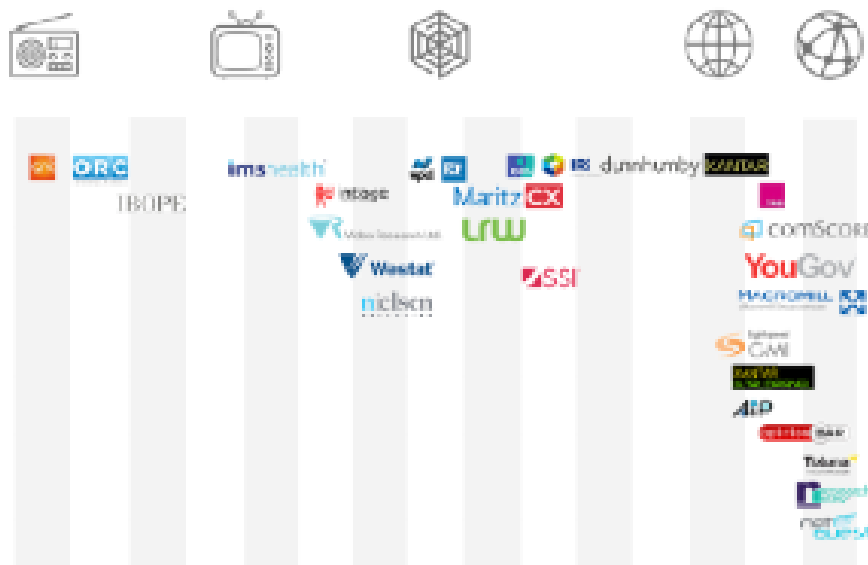
- 1. Você é um experiente profissional de pesquisa de mercado** que vem trabalhando há muitos anos com metodologias convencionais de coleta de dados, como podem ser as entrevistas pessoais e telefônicas. Você precisa realizar pesquisas e aplicar novas tecnologias para atender as demandas dos seus clientes ou porque tem necessidade de ser mais eficiente. Mesmo que você possa ter uma certa experiência em coleta de dados online, este ebook irá ajudá-lo a aprofundar na compreensão das principais diferenças com outros métodos e a fazer uma melhor escolha entre os métodos a aplicar e quando.
- 2. Você é um jovem profissional da área de pesquisa de mercado.** A coleta de dados online tem sido a sua escolha natural desde que você começou a sua carreira, mas você sente que existe uma grande diferença entre o que lhe foi explicado durante a sua formação acadêmica e a vida real. Pode ficar se perguntando como aplicar na prática tudo o que você aprendeu sobre amostragem aleatória, bases de amostragem e cotas para projetos reais. Ou então você quer saber o que mudou desde que seus estudos acabaram.
- 3. Você está atualmente estudando técnicas de pesquisa de mercado.** Você acredita que o contexto teórico a partir do qual você tem adquirido conhecimentos está longe da coleta de dados real. Ou você está tentando entender melhor como usar, num mundo cada vez mais dominado pela tecnologia, flexibilidade e velocidade, os conceitos clássicos que está aprendendo.
- 4. Você não é um pesquisador de mercado profissional,** mas precisa de coletar dados online e deseja encontrar as formas mais eficazes de fazer isso rapidamente.

Neste ebook, procuro estabelecer uma conexão importante entre a teoria e a prática de uma forma simples, breve e direta. Este ebook é simplesmente um guia. Ele não fornece detalhes completos sobre cada tópico, mas vai ajudá-lo a descobrir aquilo “que não sabemos que não sabemos”. Para mais informação, visite o nosso blog em www.netquest.com.

Como consideração final sobre o escopo deste ebook, é preciso dizer que as recomendações oferecidas aqui são aplicáveis a quase qualquer tipo de metodologia qualitativa e quantitativa de coleta de dados online. Porém, o nosso enfoque aqui serão as metodologias de pesquisa de mercado mais populares.

A internet mudou a pesquisa de mercado

DIGITAL TRANSFORMATION IS DISRUPTING MARKET RESEARCH
GLOBAL CONSUMER RESEARCH & DATA COLLECTION COMPANIES SINCE 1930



As novas tecnologias criaram novos hábitos de consumo, assim como a necessidade de estudar esses novos hábitos em seu contexto. Ao mesmo tempo, as novas tecnologias criaram novos meios para coletar dados sobre os hábitos e comportamentos do consumidor. Desde o surgimento da internet em meados dos anos 90, a pesquisa de mercado adotou gradualmente a coleta de dados online.

A cada dia que passa mais pesquisas são realizadas a partir de dispositivos conectados à internet, e novos tipos de dados são coletados através desses dispositivos. Quais são as razões para você coletar dados através da internet? Quais são os benefícios? Quais são as limitações?

A tabela abaixo permite identificar os principais prós e contras da coleta de dados online em comparação com as metodologias tradicionais.

PRINCIPAIS PRÓS E CONTRAS DA COLEÇÃO DE DADOS ONLINE

	Prós	Contras
Custo	Redução de custos devido à ausência de entrevistador	Parte da economia de custos obtida com a ausência do entrevistador deve ser usada para recompensar os participantes.
Velocidade	Mais rápido do que os métodos offline. Coleta automatizada de dados e maior número potencial de respondedores.	Projetos complexos podem ainda exigir mais tempo do que o esperado; por exemplo, se o tamanho da amostra for muito grande, é difícil de se atingir a população-alvo, ou também quando uma amostragem por cotas complexa é aplicada (ver mais adiante).
Representatividade	Irá variar de acordo com cada país em concreto, e conforme se disponha ou não de um melhor acesso a grupos sociodemográficos específicos, como podem ser os jovens e as pessoas com maior nível de instrução.	Em muitos países, a população online sub-representa alguns grupos sociodemográficos, como idosos ou de classe social baixa.
Estímulos	Quase todos os estímulos imagináveis podem ser usados na coleta de dados online, como podem ser recursos de imagem, vídeo, áudio, mapas interativos, simuladores de compra e questionários conjuntos.	Não existem contras para este tópico. Os métodos online são claramente melhores em termos de variedade de ferramentas que podem ser utilizadas na coleta de dados.
Qualidade	Devido à ausência de entrevistador, há um menor risco de respostas fraudulentas e de enviesamento induzido.	Alguns métodos de coleta de dados offline requerem a prévia consolidação e o processamento de dados para a elaboração de relatórios e para poderem ser visualizados. Nesses casos, não é possível beneficiar-se totalmente da velocidade de coleta de dados online.
Processamento de dados	Os dados de vários respondedores são simultaneamente coletados, processados e eventualmente visualizados em tempo real. É rápido e fácil.	Alguns métodos de coleta de dados offline requerem a prévia consolidação e o processamento de dados para a elaboração de relatórios e para poderem ser visualizados. Nesses casos, não é possível beneficiar-se totalmente da velocidade de coleta de dados online.



Fontes de amostras online

Qualquer processo de coleta de dados começa com a amostragem, que é o método aplicado para selecionar um subconjunto de indivíduos de uma população estatística com a finalidade de calcular as características de toda a população.

Qualquer processo de coleta de dados começa com a amostragem, que é o método aplicado para selecionar um subconjunto de indivíduos de uma população estatística com a finalidade de calcular as características de toda a população.

Existem vários métodos de amostragem. Alguns são mais adequados do que outros, dependendo dos seus objetivos de pesquisa e da sua população-alvo. Mas, antes de considerar qualquer um desses métodos, pergunte a si mesmo: onde posso encontrar participantes online para uma pesquisa online?

Quase todos os métodos de amostragem que podem ter sido ensinados a você não são aplicáveis à maioria de projetos de pesquisa. Independentemente de se você estiver aplicando métodos de coleta de dados online ou offline, a pergunta é: quantas vezes na sua vida você será capaz de usar uma amostragem aleatória pura? Esta pergunta é particularmente importante quando se utiliza a internet.

Aqui vai um primeiro conselho: esqueça-se de como você gostaria de atingir a sua amostra; pense sobre como você pode consegui-la e alinhe a sua abordagem para que ela seja realista. A amostragem online tem os seus desafios: não há censo de população da internet e você não pode interceptar pessoas aleatoriamente, como se elas estivessem andando na rua, para pedir-lhes para participar. Tente entender como a internet funciona e encontrar a melhor solução para o seu projeto de coleta de dados online, e não o contrário. Este ebook irá ajudá-lo a fazer isso.

Atualmente, quatro fontes diferentes são utilizadas por pesquisadores profissionais.

1. Painéis de acesso online

Os painéis de acesso online são comunidades online formadas por pessoas que estão dispostas a participar em projetos de pesquisa de mercado (principalmente pesquisas de opinião) em troca de algum incentivo, como pode ser dinheiro, presentes, doações ou sorteios.

Os painéis de acesso online permitem recrutar membros através de diferentes meios. Alguns são painéis abertos que permitem que as pessoas se inscrevam simplesmente acessando os respectivos sites. Os visitantes são atraídos para acessar a estes sites através de campanhas de

publicidade e acabam, assim, se convertendo em novos membros.

Outros são painéis fechados e para acessá-los é preciso de um convite individual enviado por pessoas já cadastradas em sites parceiros. Pode parecer semelhante a um painel aberto mas não é: você não pode participar diretamente em um painel fechado; vai precisar de um convite.

Hoje em dia, os painéis de acesso online são a principal fonte de participantes em pesquisas de mercado profissionais.

2. River sampling

O river sampling (“amostragem de rio”) é um método inspirado na forma como as pesquisas de opinião face a face funcionam: interceptar pessoas enquanto elas caminham pelas ruas. O método de river sampling tenta interceptar pessoas online, enquanto elas navegam na internet. Este método é aplicado mostrando anúncios em sites de alto tráfego como o Google ou o Facebook. Porém, em vez de pedir às pessoas que participem em um painel para participar em futuras pesquisas de opinião, como um painel online faz, as pessoas são convidadas a participar diretamente em uma única pesquisa de opinião online. Um incentivo direto é oferecido em troca da participação.

3. Painéis online baseados em probabilidade

É difícil construir um painel totalmente representativo da população através do recrutamento online de pessoas; formalmente falando, é impossível. Isto ocorre principalmente por duas razões: (1) em muitos países, uma parte da população não utiliza a internet, e (2) não existe um diretório completo de usuários de internet a partir do qual perfilar indivíduos aleatoriamente; como veremos mais adiante, essa limitação impede o uso da amostragem aleatória.

Os painéis baseados em probabilidade tentam superar essa dificuldade recrutando pessoas offline e oferecendo a quem não usa a internet um dispositivo de navegação e uma conexão à internet.

Desta forma, o método de amostragem probabilística clássico pode ser utilizado para a construção de um painel online. Uma vez o painel tenha sido construído, as amostras aleatórias extraídas do painel deverão também ser amostras aleatórias de toda uma população; isso irá garantir que o painel seja realmente representativo.

Os painéis online baseados em probabilidade são geralmente menores e muito mais caros do que os painéis de acesso online, sendo usados principalmente em pesquisas acadêmicas.

4. Bancos de dados proprietários

Sempre que você tem à sua disposição um banco de dados com emails ou qualquer outra forma de contatar pessoas através da internet, tem também uma boa oportunidade para realizar coletas de dados online. Dependendo do objetivo da sua pesquisa, você pode não precisar de representar uma população em geral, mas apenas o seu grupo-alvo (por exemplo, os seus clientes ou visitantes). A representatividade pode não ser um problema.

A tabela abaixo mostra os principais prós e contras de cada fonte de dados.

PRINCIPAIS PRÓS E CONTRAS DAS FONTES ONLINE DE PARTICIPANTES		
	Prós	Contras
Painel de acesso online	<ul style="list-style-type: none"> • Flexibilidade: adequado para todos os tipos de projetos de coleta de dados. • Estabilidade: o tamanho da amostra pode ser bem previsto e planejado. • Adequado para estudos de rastreamento nos quais a mesma população-alvo deve ser atingida periodicamente. • Custo moderado, mesmo para alvos difíceis de serem atingidos (se o painel estiver devidamente perfilado). • Medidas de qualidade podem ser implementadas através de uma análise comparativa da participação dos indivíduos ao longo do tempo. 	<ul style="list-style-type: none"> • Os painéis nem sempre estão disponíveis em áreas geográficas com baixo índice populacional. • Representatividade limitada: a amostragem por cotas deve ser corretamente aplicada para superar esta dificuldade.
River sampling	<ul style="list-style-type: none"> • Baixo custo. • Adequado para países de baixa população quando um painel de acesso online não está disponível. • Boa fonte de audiências online frequentes, como podem ser os adolescentes. 	<ul style="list-style-type: none"> • Problemas de qualidade: os participantes não podem ser facilmente monitorados com a finalidade de evitar identidades fraudulentas e o mau comportamento durante a participação. • Não é adequado para a realização de estudos de rastreamento, já que o tamanho da amostra é instável ao longo do tempo. • Risco de resultados tendenciosos quando todos os participantes são convidados a partir de um único site.

<p>Painéis online baseados em probabilidade</p>	<ul style="list-style-type: none"> • Boa representatividade. A amostragem por cotas não é um aspecto tão crítico como nos painéis de acesso online. • Alta estabilidade: o tamanho da amostral. 	<ul style="list-style-type: none"> • Tamanho limitado. • Alto custo. • Não disponível na maioria dos países.
<p>Bancos de dados proprietários</p>	<ul style="list-style-type: none"> • Menor custo possível: você já possui a fonte de participantes. • Dependendo de como você defina a sua população-alvo e o perfil dos participantes no seu banco de dados, a amostragem probabilística poderá ser implementada. • Os resultados podem ser complementados com informações já disponíveis (p. ex., os resultados das pesquisas de opinião realizadas com clientes podem ser comparados com os rendimentos e a segmentação de clientes). 	<ul style="list-style-type: none"> • Um banco de dados proprietário nem sempre está disponível para resolver o problema. • Adequado para coletar dados de clientes, mas não de clientes concorrentes.



Métodos de amostragem

Depois de ter em perspectiva uma visão clara das principais fontes de participantes para pesquisas de opinião, a decisão sobre os métodos de amostragem fica mais simples: o seu método de amostragem será fortemente condicionado pela fonte de participantes disponíveis para você.

Os métodos de amostragem existentes podem ser divididos em 2 categorias diferentes: métodos de amostragem probabilísticos e não probabilísticos.

Métodos de amostragem probabilísticos

Uma amostra probabilística é uma amostra em que cada unidade da população estudada tem uma chance (maior que zero) de ser selecionada para a amostra, podendo essa probabilidade ser determinada com precisão. Essas características viabilizam a produção de estimativas não imparciais dos totais de população. A amostragem probabilística sempre necessita de um quadro de amostragem: uma lista de todos os incluídos em uma população que pode ser amostrada.

A **amostragem aleatória simples** (SRS, por sua sigla em inglês) é o método de amostragem probabilística mais popular, sendo considerado um padrão-ouro: todos os métodos podem ser comparados com ele. Na amostragem aleatória simples, todos os indivíduos da população têm chances iguais de serem selecionados. Esta condição minimiza o enviesamento e simplifica a análise.

Outros métodos de amostragem probabilística são:

Amostragem estratificada: No caso de que a população apresente várias categorias diferentes ou “estratos”, cada estrato pode ser amostrado como uma subpopulação independente. Esse método (1) garante que as inferências possam ser feitas para cada subgrupo e (2) inclui as estimativas estatísticas mais eficientes, pelo menos tão boas quanto a amostragem aleatória simples, e em muitas ocasiões ainda melhores. A amostragem estratificada funciona melhor quando a variabilidade dentro dos estratos é baixa, mas é alta entre os estratos, portanto, você pode ter que adaptar o tamanho da amostra para cada estrato.

Amostras agrupadas: Quando a população pode ser dividida em grupos ou clusters e cada cluster é bastante semelhante aos outros, é possível fazer a amostragem com apenas alguns dos clusters para representar toda a população. Esta abordagem pode ser conveniente se a diminuição do número de clusters reduzir efetivamente o custo da coleta de dados (p. ex., se cada cluster for uma área de cidade dentro de uma cidade, é possível reduzir os custos enviando entrevistadores apenas para algumas das áreas de cidade). Também é possível simplificar o quadro de amostragem: no exemplo anterior, você precisa de um mapa com as áreas de cidade, e não de

uma lista de indivíduos da cidade. A amostragem de clusters pode aumentar a variabilidade das estimativas das amostras, principalmente se os clusters não forem homogêneos.

Amostragem sistemática: Este método seleciona indivíduos em intervalos regulares de uma lista ordenada da população. Por exemplo, é possível selecionar 1 de cada 50 visitantes de uma loja para ser entrevistado. Este método pode igualar ou mesmo melhorar os resultados obtidos a partir de uma amostragem aleatória simples, mas, se existir correlação na lista ordenada de candidatos a serem selecionados, as estimativas resultantes poderão ser tendenciosas

Métodos de amostragem não probabilística

A amostragem não probabilística é qualquer procedimento de amostragem que não corresponde aos requisitos de amostragem probabilística: às vezes porque alguns indivíduos da população-alvo não são acessíveis, e outras, porque a probabilidade de inclusão é desconhecida. Embora muitas pessoas a façam, a estimativa de erros não pode ser feita utilizando métodos de amostragem não probabilísticos. Os questionários telefônicos realizados a partir de números aleatórios e as entrevistas pessoais realizadas na rua em áreas de alta intensidade são exemplos de métodos de amostragem não probabilísticos.

A verdade sobre a pesquisa de mercado é que 99% da coleta de dados utiliza a amostragem não probabilística. Assim sendo, a última vez que você estudou uma estimativa de erro apresentada juntamente com o resultado de pesquisa (a chamada margem de erro e o nível de confiança), provavelmente estava perante a um mau uso disfarçado.

Felizmente, este ponto não significa que a amostragem não probabilística seja inútil. Não é perfeita, não permite calcular erros de medição, pode ser levemente ou mesmo severamente tendenciosa..., mas a informação é sempre valiosa, desde que você saiba como ela foi coletada e que, mesmo com as suas limitações, a informação atende os seus objetivos.

A **amostragem por cotas** (QS, por sua sigla em inglês) é o método de amostragem não probabilístico mais popular. Na amostragem por cotas, a população é segmentada em grupos que se excluem mutuamente, de acordo com algumas variáveis de segmentação como podem ser a idade, o sexo ou a região. A distribuição da amostra-alvo nos segmentos é feita de forma a corresponder com a população-alvo. Por exemplo, se a população-alvo for 50% do sexo masculino e 50% do sexo feminino, uma amostra de 1.000 indivíduos deve corresponder à seguinte distribuição: 500 pessoas do sexo masculino e 500 pessoas do sexo feminino.

Uma vez definido o tamanho-alvo para cada um dos segmentos ou cotas, é selecionada uma amostra não aleatória e os indivíduos são entrevistados até ser atingido o objetivo de cada cota. Ou seja, quando se tenha completado o grupo de 500 homens, se o meu próximo indivíduo na amostra for um homem, ele deverá ser descartado. Os indivíduos descartados são chamados de “cota cheia” (quota-fulls). Nesse sentido, a amostragem por cotas é semelhante à amostragem estratificada mas, em vez de amostrar pessoas aleatoriamente de cada segmento, espera-se que uma amostra não aleatória seja tão representativa quanto possível, completando as cotas-alvo por segmento.

Outros métodos de amostragem não probabilística são:

Amostragem por bola de neve: Este método começa com a identificação de um pequeno grupo de respondedores iniciais dentro da população-alvo. Estes respondedores iniciais são então utilizados para recrutar mais respondedores dentro da mesma população-alvo. É conveniente para amostras difíceis de serem conseguidas.

Amostragem de conveniência: Este nome é comumente usado para descrever qualquer outro método aplicado para obter rapidamente uma amostra não probabilística. As pessoas que se inscrevem voluntariamente para preencher formulários web e as que são fãs no Facebook são exemplos destas amostras. Você não pode calcular até que ponto essas amostras representam efetivamente a sua população, mas, mesmo assim, elas são uma amostra e podem ser úteis para o processo de coleta de dados.

Métodos de amostragem adequados para a coleta de dados online.

Como foi comentado acima, os métodos de amostragem que são adequados para a coleta de dados online dependem da fonte de respondedores que você utiliza.

Se você pretende extrair amostras de um painel de acesso online, saiba que a amostragem por cotas é a única opção realista. Os painéis de acesso online não são representativos de toda a população, mas se eles tiverem sido bem concebidos, terão a capacidade de fornecer todos os tipos de respondedores. Para garantir que uma amostra representativa seja extraída do painel, é preciso estabelecer as cotas corretas. A configuração de cotas para coleta de dados online será abordada no próximo capítulo.

A mesma premissa aplica-se também à River sampling. Neste caso, as cotas são ainda mais importantes dada a variabilidade limitada dos participantes.

Porém, isso pode variar quando se trabalha com um painel online baseado em probabilidades, já que este tipo de painel foi concebido para ser uma amostra aleatória/representativa da população.

Teoricamente, é possível confiar neste fato ao extrair amostras aleatórias do painel. No entanto, é preciso ter em mente que um painel online baseado em probabilidades ainda é uma comunidade de pessoas que são entrevistadas regularmente em troca de alguma coisa. Alguns membros do painel podem responder mais ativamente aos convites do que outros; a atividade pode ser desigual, dependendo da idade, sexo ou outro aspecto sociodemográfico. Portanto, na hora de criar uma amostra a partir do painel, esta pode ser polarizada. A amostragem por cotas é o seu melhor remédio para evitar estes desafios.

Finalmente, se você tiver acesso a um banco de dados proprietário, estará em condições de poder utilizar vários métodos de amostragem. Por exemplo, pense em uma lista completa de clientes de uma empresa: se você deseja pesquisar esse alvo específico, essa lista pode ser considerada um quadro de amostragem e, portanto, você poderá aplicar métodos de amostragem probabilística,

como a amostragem aleatória simples ou a amostragem sistemática. Naturalmente, você também poderia utilizar a amostragem não probabilística.

Às vezes, um banco de dados proprietário pode ser um quadro de amostragem ou não, dependendo de como você tenha definido a sua população-alvo. Por exemplo, imagine que você tem uma lista de clientes e clientes potenciais que se cadastraram para receber o boletim informativo de uma empresa: trata-se de um quadro de amostragem? Se você definir a sua população-alvo como “pessoas que estão interessadas em receber informações regulares da minha empresa”, então é. Se a sua população-alvo é “qualquer cliente real e potencial”, então não é. Para este segundo cenário, a amostragem por cotas é sempre uma boa opção. Porém, quando se trabalha com bancos de dados internos e, portanto, não existe um custo incremental por resposta, às vezes é melhor amostrar aleatoriamente todo o banco de dados e analisar os resultados por segmento.



Amostragem por cotas online efetivas

A amostragem por cotas é o método de amostragem padrão para a coleta de dados online.

Sempre que você encomendar um projeto de coleta de dados online a um painel online, será questionado sobre como pretende utilizar as cotas. A seguir, você encontrará as melhores práticas e recomendações sobre como definir cotas para garantir que a sua coleta de dados online seja eficaz.

1. Defina apenas cotas para variáveis que são (1) críticas para o seu projeto e/ou (2) cuja distribuição difere claramente da sua população-alvo e da população da internet.

As cotas visam superar uma possível falta de representatividade devido às diferenças observadas entre a sua fonte de dados e a população-alvo. Se você usou a amostragem por cotas para uma coleta de dados offline, deve repensar se é realmente conveniente reutilizar as mesmas cotas para as mesmas variáveis sociodemográficas. Embora você possa tê-las utilizado para configurar cotas de gênero, idade e região, elas podem não ser a aposta certa para a sua coleta de dados online.

Alguns exemplos sobre como você deve considerar as cotas destinadas a uma amostragem online:

- **Em geral, a idade é uma variável fundamental em quase todos os países.** É mais provável que os jovens sejam mais usuários da internet do que as pessoas mais idosas, exceto em certos países desenvolvidos, como a Noruega ou a Suécia, onde a utilização da internet é de aproximadamente 95%.
- **Apesar de que o uso da internet não seja muito diferente entre homens e mulheres na maioria dos países,** você deve sempre usar cotas de gênero ao elaborar amostras de painéis de acesso online. A razão disso é que os painéis precisam de ser o suficientemente grandes como para satisfazer as necessidades dos compradores de pesquisa, mas sempre considerando que as mulheres são mais pesquisadas do que os homens, já que se supõe que são elas as que tomam a maioria das decisões de compra em casa. Caso você não defina cotas de gênero, provavelmente acabará obtendo um maior número de respostas de mulheres.
- **Nos países menos desenvolvidos da América Latina, Ásia e África, a classe social é um fator crítico.** O uso da internet é muito diferente entre as classes sociais. E as classes sociais baixas comportam-se de forma radicalmente diferente das classes médias e altas. Este fator não é tão relevante nos países desenvolvidos, sendo possível evitar perfeitamente esse controle de cotas.

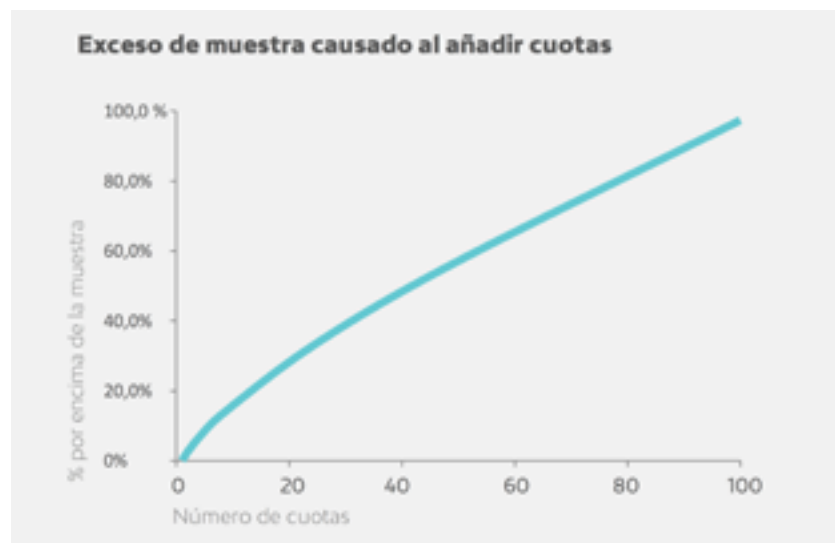
2. Menos é mais: use o número mínimo de cotas.

À medida que as cotas aumentam a complexidade do processo de coleta de dados, elas afetam o cronograma do projeto e os custos derivados da coleta de dados. Esse problema é inerente à própria amostragem por cotas, mas cabe destacá-lo aqui para quando você pensar em efetuar uma coleta de dados online.

Reflita sobre as principais diferenças entre a coleta de dados online e offline. Quando você realiza uma pesquisa do tipo face a face tendo definido uma cota de gênero, uma vez que o seu alvo de população masculina foi atingido, você para imediatamente de interceptar homens na rua, porque sabe se alguém é homem ou mulher por simples observação!

Mas o processo online é totalmente diferente: você envia um lote de convites por email para um conjunto de candidatos, então espera até que as pessoas recebam e abram os emails, acessem o questionário online e completem a sua participação. Quando o seu alvo de população masculina foi atingido, você não pode impedir que outros candidatos masculinos acessem o questionário, porque não pode suprimir um convite enviado por email após o mesmo ter sido entregue. Precisarás descartar essas participações de excesso de cota, já que elas não têm nenhum valor para você, embora tenham um custo para o painel de acesso online.

Quanto mais complexas forem as suas cotas, haverá um maior custo e mais tempo vai ser necessário para completar a coleta de dados online. O gráfico abaixo mostra o número aproximado de pessoas que talvez você precise adicionar ao tamanho da amostra devido a um excesso de cotas, assim como para um tamanho de amostra de 1.000 e para cotas com tamanho uniforme. Por meio deste gráfico, o estabelecimento de uma quota com tamanho = 2 (dividindo o tamanho-alvo de 1.000 em dois grupos-alvo de 500 indivíduos) exige 2,5% mais de respondedores. Uma cota com um tamanho = 8 (8 grupos-alvo de 125 indivíduos) aumenta a amostra necessária em 13,1%. Por último, dividindo o nosso tamanho-alvo em 100 cotas, teremos que praticamente dobrar o tamanho da amostra (+ 96,6%).



3. Decida sabiamente se é preciso interligar cotas ou não

Falamos de cotas não interligadas quando as metas em cada cota são definidas independentemente. Por exemplo, para um tamanho de amostra de 1.000 respondedores, você pode requerer 500 (50%) homens e 500 (50%) mulheres, e 500 (50%) jovens e 500 (50%) pessoas mais velhas. Ao fazer isso, você corre o risco de acabar com 500 homens jovens e 500 mulheres mais velhas. Este resultado é perfeitamente compatível com a sua definição de cota, mas, você realmente gosta dessa distribuição?

Para evitar esse efeito, você pode utilizar cotas de interligação, ou seja, definir um tamanho de destino para cada combinação de variáveis: 250 (50% x 50%) homens jovens, 250 (50% x 50%) mulheres jovens, 250 (50% x 50%) homens mais velhos e 250 (50% x 50%) mulheres mais velhas.

Você deve usar cotas interligadas ou não interligadas para coletar dados online? Isso irá depender das suas prioridades. As cotas não interligadas supostamente simplificam a coleta de dados: uma definição de cota interligada sempre corresponde à definição equivalente de cota não interligada, mas não vice-versa. Você pode constatar facilmente esse fato revisando o exemplo acima.

Em suma: o uso de cotas não interligadas costuma ser mais barato e mais rápido. Mas isso nem sempre é verdade. No caso do exemplo acima, se o painel de acesso online for severamente distorcido em relação aos jovens e aos homens, é muito provável que, quando você complete as suas primeiras 500 pesquisas, a grande maioria delas procedam de homens jovens. Isso significa que o restante das pesquisas pendentes (500), precisará proceder do seu alvo mais difícil de ser atingido: mulheres mais velhas. Às vezes é melhor interligar cotas para permitir um progresso equilibrado em todos os alvos ao mesmo tempo.

O que você pode fazer para superar esse desafio? O nosso conselho é: (1) permitir cotas não interligadas somente se o painel de acesso online estiver uniformemente distribuído entre alvos (p. ex., em vez de ter uma distribuição de gênero 50% por 50%, ter uma de 45% por 55%, mas não de 30% por 70%), e (2) utilizar sempre cotas interligadas quando você precisar de analisar alguns sub-alvos com uma certa profundidade (p. ex., se você quiser analisar as diferenças entre homens idosos e mulheres idosas).

Planejando a coleta de dados online

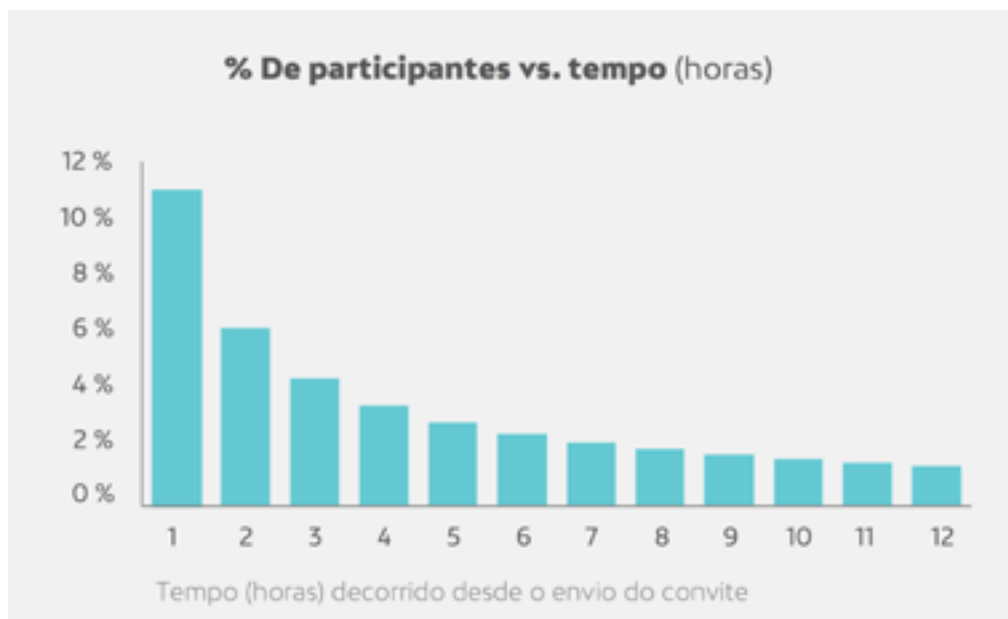
A velocidade é uma clara vantagem da coleta de dados online.

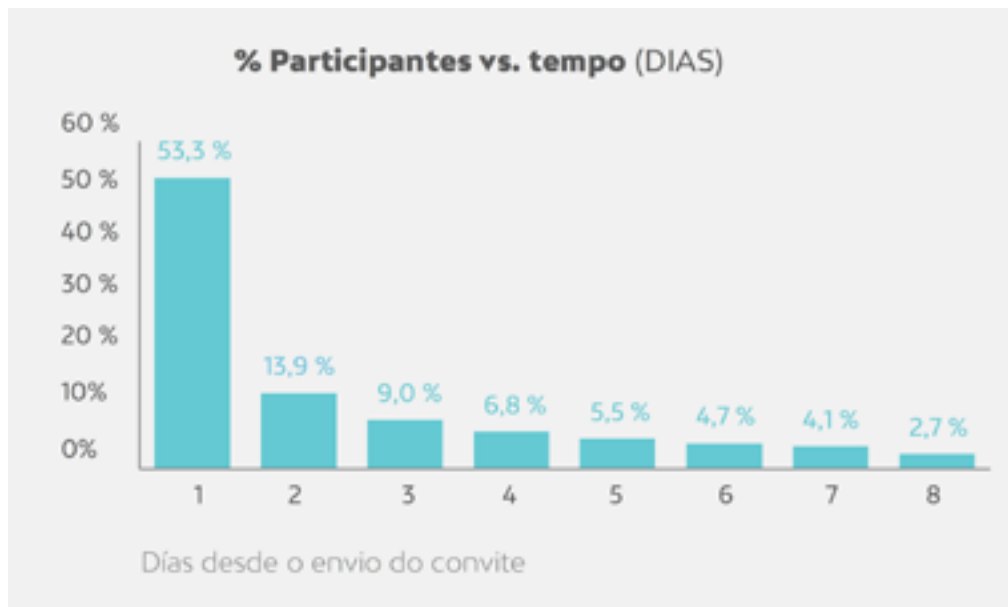
É muito mais rápido enviar milhares de convites por email e permitir que participantes simultâneos completem pesquisas do que telefonar para pessoas através de um call center ou interceptar pessoas que caminham pelas ruas.

Os trabalhos de campo online normalmente demoram entre 2 e 5 dias, dependendo do tamanho da amostra, da população-alvo e, é claro, do painel de acesso online utilizado por você. Em princípio, a coleta de dados online pode ser feita mais rapidamente do que no prazo de 2 dias. Neste caso, você vai precisar convidar mais candidatos para participar ao mesmo tempo. Mas isso pode não ser uma boa ideia para os seus dados.

Considere quem são mais propensos a participar em pesquisas online quando for estabelecer uma janela limitada para participar: usuários de internet pesados, viciados em smartphones ou alguns funcionários de escritório. Às vezes, os pesquisadores estão realmente preocupados com a representatividade das fontes de participantes. Porém, o impacto de um mau procedimento de coleta de dados é muitas vezes subestimado.

Os gráficos abaixo, elaborados a partir dos dados dos 6 maiores painéis da Netquest, ajudam a compreender o impacto da limitação do tempo de participação disponível.





A participação em pesquisas online diminui com o tempo: 11% das participações são completadas dentro da primeira hora, e 22% nas primeiras três horas. Mas se você fechar o seu trabalho de campo em apenas 3 horas, estará perdendo 78% das participações, tornando a sua amostra mais cara. E o que é ainda pior: você não sabe se os participantes que não têm permissão para participar são diferentes daqueles que participaram, influenciando assim o resultado da sua amostra.

Em resumo, a coleta de dados online é mais rápida do que a coleta de dados offline, e pode ser ainda mais rápida, mas com o risco de predispor a sua amostra a um enviesamento, aumentando os custos.



Pesquisas eficazes para o trabalho de campo online

Se você tem alguma experiência em pesquisa de mercado quantitativa, provavelmente conhece os fundamentos da elaboração de um questionário.

Os fundamentos sempre serão os mesmos, não importa se você pretende coletar dados online ou offline. A norma procedimental de elaboração de pesquisas é muito extensa para abrangê-la totalmente aqui, mas como um lembrete simples, a sequência apropriada para fazer este trabalho é a seguinte:

1. Definir o objetivo principal do questionário: qual é o objetivo da pesquisa?
2. Determinar as perguntas a serem feitas.
3. Selecionar o tipo de pergunta mais adequado para cada pergunta, entre diferentes opções (perguntas abertas ou fechadas, resposta única ou múltipla, perguntas de classificação, perguntas de avaliação, etc.).
4. Elaboração do texto adequado para cada pergunta, de acordo com a população-alvo.
5. Determinação da sequência de perguntas e desenho do layout do questionário de maneira que tudo faça sentido para o participante.

As melhores práticas para elaborar questionários offline eficazes são também válidas para os questionários online:

- Seja conciso, breve e claro. Evite frases complexas.
- Evite pedir duas coisas diferentes na mesma pergunta.
- Evite perguntas negativas que possam confundir os participantes.
- Peça respostas precisas.
- Ao utilizar perguntas fechadas, certifique-se de que você está oferecendo todas as respostas possíveis (considere oferecer a opção “outros”).
- Evite perguntas sofisticadas.

No entanto, existem algumas recomendações específicas para pesquisas de opinião online.

1. Use várias páginas em vez de uma rolagem interminável.
2. Pergunte uma coisa de cada vez e limite o questionário a uma pergunta por página. Não permita que o respondedor navegue pela página escolhendo a ordem de resposta de cada pergunta.
3. Estabeleça uma duração moderada para a pesquisa de opinião. Como regra geral, 20 minutos, no máximo; 5 minutos se os seus respondedores foram instados a utilizar os seus dispositivos móveis.

4. Utilize questionários online responsivos que se adaptem adequadamente a qualquer tamanho de tela normalmente existente (PC, tablet, celular).
5. Seja específico na sua forma de dizer as coisas. As pessoas procuram textos online por palavras-chave e mensagens-chave.
6. Use imagens e ícones sempre que possível, já que isso permite uma melhor experiência de usuário para os participantes.
7. Inserir vídeos em um questionário é uma possibilidade exclusiva das pesquisas de opinião online que marca a diferença com as pesquisas offline. Você pode se beneficiar com isso (p. ex., testes criativos ou comerciais), mas considere que a participação será reduzida devido a vários fatores: compatibilidade do vídeo com navegadores antigos, conexões de banda larga limitadas a partir de dispositivos móveis e, em alguns casos, as pessoas não estarão dispostas a ativar a função de áudio dos seus dispositivos (p. ex., se estiverem viajando em um transporte público).
8. Cumprir os padrões da Internet, não “reinventar a roda”. As pessoas esperam que, ao clicar em um botão de rádio, os outros botões de rádio próximos fiquem desmarcados. Cada elemento de uma página web (caixas de seleção, caixas de entrada de texto, botões de envio, texto sublinhado, etc.) deve se comportar como é esperado.
9. Evite baterias de perguntas. De fato, as pesquisas têm demonstrado que os benefícios potenciais de agrupar perguntas com uma escala de resposta partilhada (tempo de resposta mais rápido e maior correlação inter-item) não compensam os inconvenientes (pior experiência de usuário, maior taxa de abandono e menor diferenciação da resposta).
(*)
10. Aproveite as possibilidades oferecidas pela tecnologia: simulação de compra, questionários conjuntos online, recursos de voz, reconhecimento facial, análise da movimentação dos olhos, etc. Há uma vasta quantidade de novas e promissoras funcionalidades disponíveis que estão atualmente sendo utilizadas na coleta de dados online.

(*)Melanie Revilla, Carlos Ochoa. *Telematics and Informatics*, abril 2016. “An experiment comparing grids and item-by-item formats in web surveys completed through PCs and smartphones”

Dados de qualidade

Se você está considerando a possibilidade de usar a coleta de dados online pela primeira vez, ou mesmo que já tenha usado essa metodologia várias vezes, poderá se preocupar com a seguinte pergunta: até que ponto a qualidade dos dados online é comparável com a dos dados offline?

Existem alguns aspectos que trabalham em contra da qualidade na coleta de dados online mas, ao mesmo tempo, existem muitos outros que promovem a qualidade. Na verdade, existem várias maneiras de controlar a qualidade na coleta de dados online de uma forma que você possa superar os principais inconvenientes e, ao mesmo tempo, beneficiar-se das virtudes desta metodologia.

PRÓS E CONTRAS DA COLETA DE DADOS ONLINE EM TERMOS DE QUALIDADE

Prós	Contras
<ul style="list-style-type: none">• Ausência de enviesamento por parte do entrevistador: as respostas não são distorcidas devido a diferenças no estilo social e personalidade dos entrevistadores ou à apresentação de questões específicas.• Confiabilidade melhorada ao fazer perguntas sensíveis (preferências políticas, estado de saúde, religião, etc.) devido ao anonimato do respondedor. (**)• Nenhum erro de transcrição. Os participantes fornecem respostas diretamente no questionário.• Os metadados são facilmente coletados para avaliar a qualidade, permitindo a tomada de ações corretivas. Dados como podem ser o tempo de participação, tempo por questão ou número de interrupções do questionário são coletados por quase todos os softwares de pesquisa profissionais.	<ul style="list-style-type: none">• Manter um controle rigoroso sobre quem está participando é difícil (p. ex., em entrevistas face a face, é fácil observar visualmente a idade aproximada e o sexo dos participantes, mas na internet não).• Maior risco de satisficing (decidir e agir satisfazendo um mínimo de requisitos para atingir um objetivo) devido à ausência de entrevistador e ao uso de incentivos. O satisficing é uma estratégia que os respondedores podem usar para responder as perguntas com o mínimo esforço exigido. (***)• A comparabilidade dos dados pode ser afetada pela alta diversidade de dispositivos utilizados pelos respondedores para participar: diferentes navegadores, versões de navegadores, dispositivos, tamanhos de tela, etc.• Ao recompensar a participação, o comportamento do participante pode ser afetado. Por exemplo, um painel que oferece apenas um tipo de incentivo em troca da participação pode sobrerrepresentar uma determinada população-alvo. Por outro lado, se o painel somente dá incentivos aos participantes que correspondem à população-alvo (p. ex., para aqueles que passam as perguntas de filtro), as pessoas podem ser encorajadas a mentir.

(**) <http://web.comhem.se/u22779327/16/k16docs081211.pdf>

(***) <http://www.businessdictionary.com/definition/satisficing.html>

Considerando o resumo apresentado acima, é óbvio que a coleta de dados online oferece benefícios e desvantagens. Mas isso ocorre igualmente com qualquer outra metodologia. No entanto, a coleta de dados online fornece uma vantagem crítica em comparação com qualquer outra solução: a maioria dos inconvenientes pode ser reduzida (ou mesmo eliminada) através da seleção de uma fonte de dados confiável e da introdução de controles de qualidade no questionário online.

7 estratégias comuns para melhorar a qualidade são:

- 1.** Compare dados da pesquisa com dados do painel para detectar possíveis identidades fraudulentas. Por exemplo, ao perguntar a idade e o sexo, é fácil validar a resposta dada pelo painalista a partir da informação fornecida quando ele se cadastrou no painel.
- 2.** Ponderar a atenção prestada pelos participantes à pesquisa, comparando respostas dentro da pesquisa para avaliá-la. Você pode fazer a mesma pergunta duas vezes (de uma maneira levemente distinta) num ponto diferente do questionário.
- 3.** Faça uma verificação de manipulação instrucional (IMC, por sua sigla em inglês) quando a duração do questionário for longa (> 15 minutos). As verificações de manipulação instrucional são perguntas que testam se o participante está prestando realmente atenção ao questionário, como, por exemplo: "Se você está lendo esta pergunta, por favor, selecione a terceira opção de resposta".
- 4.** Recompense as pessoas em troca da sua participação, não em troca de igualar a população-alvo que você deseja pesquisar. Se você paga somente àqueles que respondem "sim" a sua pergunta de filtro (como, p. ex., "Você usa este produto?" "Você conhece essa marca?") você vai acabar encorajando as pessoas a mentir na próxima pesquisa que enviar a elas.
- 5.** Defina um tempo mínimo necessário para concluir a pesquisa e descarte os participantes que não o atingem. É uma maneira fácil de detectar pessoas que estão respondendo sem prestar atenção.
- 6.** Verifique as respostas dadas às perguntas abertas para detectar respostas erradas (p. ex. "---", "aaaaa"), e suprima a participação completada como um todo, não apenas a resposta. Se alguém tiver completado uma pergunta aberta desta forma, provavelmente ele ou ela respondeu indevidamente toda a pesquisa.
- 7.** Coletar tantos metadados (dados relacionados com como e onde os dados foram coletados) quantos você puder, poderá ajudá-lo a identificar problemas. Essa é uma tarefa realmente simples quando se realizam pesquisas de opinião online. Por exemplo, alguns respondedores podem ter tido problemas para visualizar uma pergunta com uma determinada combinação de navegador-dispositivo. Se você coletou esses metadados, você será mais propenso a entender o problema.

Se você está considerando a possibilidade de usar um painel de acesso online, vale a pena entender as informações apresentadas abaixo para garantir que você irá escolher o caminho certo para o seu projeto:

1. Como o painel recruta novos membros? É um painel aberto ou fechado? Onde os membros são recrutados? O cadastro é restringido com a finalidade de reduzir a autoseleção?
2. Como o painel recompensa a participação? Todas as participações são recompensadas ou apenas aquelas que terminam em uma pesquisa de opinião válida? Os painéis que somente pagam por pesquisas de opinião válidas são mais susceptíveis de sofrer respostas fraudulentas em estudos realizados com populações-alvo difíceis de atingir.
3. Quantas pesquisas de opinião os painelistas completam por mês? Alguns painéis enviam uma ou duas pesquisas de opinião por dia aos seus membros, como se quisessem “profissionalizá-los”.
4. Qual é o tamanho real com base na entrega máxima do painel? Este é um fator-chave que diz muito sobre um painel. A menos que um painel esteja em fase de criação, uma empresa de painel que gera uma quantidade substancial de negócios com um pequeno tamanho máximo de amostra, provavelmente está mais próxima de ser uma empresa de river sampling do que um verdadeiro painel. Para avaliar isso, você deve solicitar informações sobre a entrega máxima, que é o tamanho máximo de amostra que o painel pode oferecer de uma vez para a realização de uma pesquisa de opinião. Isso permite obter uma informação muito mais relevante do que simplesmente querer saber o tamanho do painel: muitas empresas dispõem, como painelistas, de pessoas que se cadastraram no painel há muito tempo, mas que já não estão mais participando ativamente.
5. O painel está certificado com uma norma de qualidade como a ISO 26362? As normas ISO garantem que os painéis foram auditados e estão dispostos a partilhar de forma transparente as informações indicadas acima.



Privacidade

A privacidade está se tornando uma das grandes preocupações das pessoas que utilizam a internet.

A pesquisa de mercado online é diretamente afetada por este fenômeno. Em última análise, os pesquisadores de mercado coletam dados pessoais e certos dados coletados podem ser percebidos como sendo altamente sensíveis por parte dos participantes.

Como pesquisador de mercado, esta é uma preocupação muito importante. Os pesquisadores de mercado estão obrigados a trabalhar conforme as disposições do código de conduta da indústria (p. ex., o da Sociedade Europeia para a Pesquisa Social e a Pesquisa de Mercado, ESOMAR), e a cumprir todas as obrigações legais que lhes são atribuídas conforme o lugar onde é feita a coleta de dados.

Você deve seguir as recomendações acima, principalmente por duas razões:

- É o seu dever como profissional de pesquisa de mercado. Devemos criar confiança entre a população se quisermos preservar a nossa profissão.
- Os dados podem ser gravemente afetados se você não proporcionar tranquilidade aos participantes: somente aqueles que não estão preocupados com a privacidade poderão querer participar ou, o que é ainda pior, as pessoas podem acabar fornecendo informações falsas.

Siga estas recomendações sempre que for coletar dados online:

- 1.** Forneça acesso a uma Política de Privacidade, isto é, uma declaração onde você pormenoriza a sua política de coleta de dados e a liberação de informações sobre os participantes. Se você estiver usando um painel de acesso online, esta responsabilidade será do próprio painel.
- 2.** É preciso informar o painalista, consistentemente e de forma transparente, ao longo de todo o seu período de colaboração. Não se conforme apenas em apresentar a sua Política de Privacidade quando for coletar dados particularmente sensíveis: você também deverá incluir na sua pesquisa de opinião mensagens claras em locais-chave que façam referência a isso.
- 3.** Você deve considerar, como casos especiais e que requerem um gerenciamento diferente, as seguintes ações: instalar/coletar cookies de terceiros, coletar informações de navegação, solicitar dados de localização geográfica, solicitar endereços postais ou de email e pedir números de telefone.
- 4.** Você deve cumprir as disposições legais do país onde se realiza a coleta de dados, não só as do seu próprio país. As referidas disposições legais podem variar bastante de um país para o outro.

5. Leve em consideração tanto as diferenças entre os países como a respectiva preocupação com a privacidade. Por exemplo, nos Estados Unidos e na Europa é perfeitamente plausível solicitar dados sobre a renda familiar, o banco ou o número de filhos; mas na América Latina isso pode ser percebido como fora de lugar devido ao risco de sequestros.

Finalmente, para melhor ajudá-lo a avaliar até que ponto as pessoas estão dispostas a participar em projetos de coleta de dados não convencionais (alguns deles intimamente relacionados com a privacidade), os resultados de uma pesquisa de opinião experimental realizada pelo painel da Netquest aparecem indicados abaixo:

ACTIVIDAD DISTINTA A ENCUESTAS	VOLUNTAD DE PARTICIPAR
Testar um produto em casa	72,4 %
Enviar fotos realizando algumas atividades	54,9 %
Digitalizar códigos de barras de produtos	51,8 %
Compartilhar a sua localização GPS	20,0 %
Registrar o seu rosto para medir emoções enquanto você assiste um vídeo	17,4 %
Instalar um medidor para rastrear informações de navegação no seu smartphone	17,0 %
Instalar um medidor para rastrear informações de navegação no seu computador	15,8 %



Relatórios e visualização

Os relatórios não fazem parte do processo de coleta de dados. A visualização (comunicação visual de dados) também não. Ambas as atividades são o ponto de partida da análise dos resultados.

Porém, a pesquisa online alterou levemente esses aspectos: os dados são coletados em tempo real e os clientes também querem ver os resultados em tempo real. Isso requer que as empresas de pesquisa de mercado ofereçam painéis de controle online, ferramentas de relatório do tipo “faça você mesmo”, e infográficos pré-construídos que são preenchidos com os dados assim que estes chegam.

Naturalmente, estas soluções extremamente rápidas limitam o toque humano. O alto valor fornecido por um analista especializado e totalmente mergulhado nos dados na tentativa de encontrar insights surpreendentes, não se encaixa muito bem com os resultados em tempo real. Mas a velocidade é o mantra das empresas modernas, e você, como pesquisador de mercado, não deve lutar contra ela.

O pensamento acima não significa que não há valor na análise de resultados individualizada e profunda, mas sim que você deve sabiamente escolher entre as duas abordagens, dependendo do cliente e do problema de pesquisa; ou mesmo se é preciso combiná-las. Por exemplo, clientes online (como podem ser sites de comércio eletrônico) estão realmente acostumados a lidar com dados, então se sentirão confortáveis trabalhando com um painel online; enquanto os clientes mais tradicionais poderão não estar dispostos a fazer isso sem o apoio de recomendações personalizadas.

A maioria das soluções de software utilizadas para coletar dados online incluem algum tipo de relatório online. Além desses relatórios básicos, existem soluções especializadas para implementar relatórios avançados e a sua visualização. A título de exemplo, abaixo listamos alguns deles:

- Dapresy (www.dapresy.com): Fornece ferramentas de visualização de dados e ferramentas de business intelligence (inteligência empresarial) que permitem aos pesquisadores processar, analisar e construir painéis baseados em infográficos.
- Tableau (www.tableau.com): Ajuda os pesquisadores de mercado a ver e entender melhor os seus dados coletados. Conecte-se a quase qualquer banco de dados, arraste e solte para criar visualizações e partilhe-as com um simples clique.
- Qlik (www.qlik.com): Oferece soluções de plataforma intuitivas para visualização de dados de autoatendimento, aplicativos de análise orientados, análises incorporadas e relatórios.

- Microsoft BI (powerbi.microsoft.com): Conjunto de ferramentas de análise de negócios para analisar dados e partilhar insights. O MBI permite que você monitore a sua empresa e obtenha respostas rapidamente com painéis de controle avançados disponíveis em todos os dispositivos.
- IBM Cognos (www-03.ibm.com/software/products/en/cognos-analytics): Uma plataforma de BI corporativa para descoberta de dados regulamentados e relatórios gerenciados que automatiza a criação de relatórios e painéis para que os usuários tenham a liberdade de fazer isso por conta própria.
- Microstrategy (www.microstrategy.com): Oferece relatórios corporativos e descoberta de dados para produtividade móvel e telemetria em tempo real, e disponibiliza às organizações toda a gama de recursos para transformar os seus dados em inteligência do mundo real.

Contato

USA & Canadá

americas@netquest.com

Portugal

portugal@netquest.com

Espanha

iberia@netquest.com

Europa & Ásia

rfqinternational@netquest.com

netquest.com

Colômbia

colombia@netquest.com

Brasil

atendimento@netquest.com

Cone Sul

cs.sales@netquest.com

México

latamnorte@netquest.com

genuine data