FIFTH EDITION

THE PRACTICE OF MARKET RESEARCH

From Data to Insight

Pearson

Yvonne McGivern

The Practice of Market Research

Industry Insight 8.3

Plant power

Australia is one of the world's largest meat-eating populations. According to 2018 data from the OECD, an average Australian consumes close to 95kg of meat per year. However, for health and ethical reasons – as well as the rise of 'flexitarianism' – meat consumption habits are changing across the country. Mintel research shows that almost one-fifth of urban Australians avoided, or were intending to avoid, red meat in 2017 and one half of them stated that they believed it is healthier to do so. Consumer movement into plant-based foods aligns with Mintel's 2017 global food and drink trend, 'Power to the Plants', which describes how people's preference for natural, simple and flexible diets is driving expansion of vegetarian, vegan and other plant-focused innovation across multiple product categories. When it comes to product innovation, Mintel's Global New Product Database shows that food and drink launches featuring the 'vegan/no animal ingredients' claim accounted for 2 per cent of all Australian food and drink launches in 2013; in 2017 it has risen to 5 per cent. According to Mintel research, more than one in three urban Australians prefer to buy products that are produced using sustainable sourcing methods, while almost one third are willing to pay a premium for everyday goods that are safe to use – for example, have no additives, or do no harm to the body. This is also reflected in Mintel's trend 'Buydeology', which underscores the need for meat-alternative manufacturers to pay close attention to ingredient purity – as well as to sourcing ethics – when catering to consumers who are increasingly basing their purchasing decisions on a company's ethical stance.

Source: Adapted from McMillan, S. (2018) 'Plant power Down Under', *Impact*, 22, pp. 12–13. Used with permission.

Box 8.5

Examples: sources of academic and scholarly material

- Jisc Library Hub Discover which offers access to details of materials held in many UK and Ireland national, academic and specialist libraries (https://discover.libraryhub.jisc.ac.uk/)
- Directory of Open Access Journals (https://doaj.org)
- Google Scholar (https://scholar.google.com)
- Microsoft Academic (https://academic.microsoft.com)
- Bielefeld Academic Search Engine or BASE (https://www.base-search.net)
- CORE (https://core.ac.uk)
- Semantic Scholar (https://www.semanticscholar.org)
- Baidu Scholar (https://www.baidu.com).

Consumer-generated content or data

Consumer-generated content – that is, material published in online discussion forums, blogs, microblogs, content communities, virtual worlds and other social media – is in

effect published data. This type of material is also known as user-generated content (UGC) or user-created content, or volunteered data (Kitchin, 2014). It is content produced by consumers and users that is publicly available. It can be in any format – text, audio, images or video. It gives marketers and other interested parties direct access to the consumer. It contains information about the users who generated it, including locations and networks of contacts, as well as attitudes, opinions and preferences that relate to organisations, products and services. For researchers it can be a less expensive way of getting insight than traditional qualitative or quantitative research. Vriens et al. (2019) describe it is a 'valuable and viable source of consumer insights'. They note its benefits as an alternative to surveys–the fact that the data are time-stamped (allowing the evaluation of the impact of a campaign or event); that it is 'easy to zoom in and out'; and that it facilitates the study of a large set of brands. There are ethical, legal and regulatory issues, of course, and it may be that these eventually restrict the use of user-generated content for research, and with the use of blockers consumers themselves and the browsers they use may also restrict data collection (Nunan, 2020). The use of UGC also has other limitations, including lack of wider context to enable understanding of the reason for the opinion or preference and the ability to distinguish genuine consumers from those pretending to be genuine consumers.

Many organisations use these data (where ethically and legally appropriate) for research purposes. The type of research includes social media monitoring or social media listening (from automated monitoring to ad hoc research), social media mining, crawling, extracting, gathering or 'scraping' content (manually or using automated processes) from sites and platforms, mining it and analysing it. When the aim is to find out what people are saying about your company or your brand, say, or a topic relevant to your organisation, it is sometimes referred to as 'buzz monitoring' and the search for and analysis of those conversations as 'buzz mining' and opinion mining and sentiment analysis. It is useful in identifying issues that are being talked about (the content), perceptions consumers hold, reactions to events, and views and opinions being expressed and the way in which they are expressed. It is useful for uncovering issues or concerns about a product or service, for getting an idea of a brand's image or reputation, for tracking the impact of advertising and other marketing activity and for gathering intelligence about competitors. Data mining and data analytics techniques including machine learning can be used to identify patterns and to cluster and segment consumers.

Organisations can gather consumer-generated content themselves. For example, there are platforms that enable you to access and use public Twitter data; to gather tweets from Twitter users; to download your own tweets; to track uses of key words or hashtags in real time across Twitter; to connect to Twitter data streams to gather data; and to search past tweets. Alternatively, you can buy data from authorised data brokers or resellers.

Industry Insight 8.4 gives an example of the use of social media research. Mobile network Three and researchers at Mindshare UK analysed social media data to get a more rounded view of Three's target audience. The findings were used to make changes to website content, and to modify brand positioning and pricing strategy.

Industry Insight 8.4

What do you think of Three?

The brief

Mobile network operator 'Three' wanted to understand better a segment of the market where it had previously underperformed versus its competitors. Three had formed a hypothesis that its brand had a reputation for being 'cheap for a reason' amongst this target segment and that this was limiting customer acquisitions. However, this attitude was not apparent in the findings from brand trackers and other 'traditional' forms of measurement. So, on the basis of this hypothesis, the company put the following three key questions to the research agency, Mindshare UK:

- How do people talk about Three?
- What is Three associated with?
- What are the kinds of questions associated with Three?

The insights generated by the research would be used to inform the next phase of research and strategies relating to target customer, brand, pricing and hardware.

The research

Mindshare suggested using a combination of social listening and search data to uncover consumer attitudes towards the Three brand that the surveys weren't revealing. Due to its large scale (high number of mentions), and unprompted nature, analysis of social conversation can reveal deeply held consumer attitudes and biases in a way that questionnaires and traditional brand trackers have been unable to do. Tools used for this analysis included Brandwatch, Crimson Hexagon, Hitwise and some bespoke Mindshare analytics tools.

Mindshare gathered a dataset of brand mentions of 'Three' stretching back five years. First, they analysed the mentions for sentiment, which gave a broad indication of general attitudes towards the brand. Next, they layered on emotional analysis and segmentation. After that they analysed the mentions to examine the vocabulary used to describe brands. Finally, they segmented the conversation into six key themes: 'Network Quality', 'Data', 'Upgrades', 'Roaming', 'Customer Service' and 'Pricing'. This enabled the researchers to understand which topics were most frequently discussed in relation to Three. To provide context to the analysis, they did the same for Three's key competitive set – O2, Vodafone and EE. The analysis provided extensive insight into the key hypothesis.

However, the agency felt that the question merited further exploration and corroboration and so they decided to look at search behaviour. They wanted to add to the social findings by digging deeper and analysing behaviour in relation to the brand at a layer below active conversation. They did this using Google data. They mapped associated searches to Three and competitors – these are searches made in the same session as brand searches. Mapping these associations revealed the mindset of the audience searching for the brand. They also examined website traffic data around Three and competitors. First, they looked at key websites visited when 'Three' or competitor brands were searched for. Second, they analysed upstream and downstream traffic from the Three and competitor website homepages, to gain insight into the consumer journey. The obvious challenge came from creating a Boolean search string that effectively scraped mentions of 'Three' the brand, without including irrelevant content. They addressed this in two ways:

- Accuracy: They constructed a complex query string that pulled in relevant 'Three' brand mentions, excluding irrelevant terms, and adding in qualifying terms such as 'network', handsets etc.
- Consistency: When gathering datasets for competitors with less ambiguous names, they used the same approach so that they could make a fair comparison between brand conversations. To do this they made sure that limitations around searches for 'Three', such as limiting brand mentions to those also including handset or network comments, were applied across the board to all competitors.

Use of social data in research runs a risk of presenting a polarised view – with people actively posting on social media when particularly angry or happy. The steps the agency took to mitigate this included:

- Benchmarking: It was imperative to compare Three's results to those of competitor brands. Any risk of exaggerated results is avoided by a shift in emphasis between absolute results, to comparative performance between brands.
- Inclusion of other datasets: The agency takes the view that the use of multiple data sources and research techniques yields the best results. For this project they felt that combining the analysis of active conversation with more passive search behaviour would approach the brief questions from different, complementary angles of consumer behaviour.

By segmenting conversation into topic areas, and performing more complex linguistic analysis, the agency gained richer and more actionable insights. Corroborating social findings with search behaviour worked well to provide a more holistic view of the audience – a combination of opinion and behaviour. This was particularly useful for Three as it allowed them to show how certain brand perceptions translate into different behaviour for Three vs. its competitors.

The findings

The findings did not prove the 'cheapness' hypothesis as expected: negative pricing-related conversation was no more prominent for Three than for competitors. This was validated in a subsequent phase of quantitative research where no strong implicit or explicit associations of 'cheapness' were attributed to Three by its target segment. The research showed that Three is less discussed than competitors, and overall less associated with either negative or positive adjectives. Three was less considered than competitor brands, driving less strong feelings. This indicated lack of brand loyalty was corroborated with upstream and downstream website traffic, which contained a much higher proportion of independent forums and price comparison websites than for EE and O2, where customers went straight to these pages and didn't compare operators.

The business impact

The research validated some of the Three team's long-held suspicions about the brand's reputation, showing how attitudes people have towards a brand manifest themselves clearly in social and search behaviour. The findings gave the team proof points around its brand reputation compared to competitors, and this helped push through change.

Source: Adapted from Mindshare UK and Three, 'Three brand reputation analysis: Using social conversations and search insight to uncover implicit attitudes towards Three Mobile Network', Winner, MRS Awards 2018. Used with permission.

External sources

- Websites
- Statistical services
- Data services
- Data portals and hubs
- Archives, repositories and libraries
- Social media

Internal sources

- Data lakes, databases, data warehouses
- Insight management or business intelligence system
- CRM system
- Enterprise intelligence system

Figure 8.1 Where to look for existing data

Internal sources

With the continued rise in the number and use of digital devices, the integration and networking of devices, including the Internet of Things (IoT), the increase in computing power and the availability of vast amounts of relatively cheap storage, lots of data are not only generated but they are processed, stored and analysed by organisations of all types. The data collected are used in core functions to understand and manage the operation of the organisation, its products and/or services, and its users and stakeholders, in an efficient and effective way. A lot of these data are the sort that marketers and advertisers, in particular, find useful.

A lot of these data are gathered automatically without the user necessarily knowing that data gathering is happening; and some data are volunteered, including personal data when registering for a loyalty card or as a customer or member on a website or downloading apps. The data collected are aggregated, stored and processed in data lakes, databases, data warehouses or data management platforms (DMP) from which they can be retrieved for analysis. These storage and retrieval systems can be designed to function as enterprise intelligence systems (EIS), management information systems (MIS) or business intelligence systems (BIS) for planning and control, strategy development, for example. They can be set up to serve more specific functions such as marketing (marketing information systems or MkIS) or customer relationship management (CRM systems) or research and insight management. Before we move on to look at these storage and retrieval systems it is worth looking at how some of the data particularly relevant in a market research context are gathered from customer transactions and interactions.

Data from loyalty cards and customer databases

We mentioned 'loyalty' cards above. Loyalty cards, sometimes called 'reward' or 'club' cards, are used by many companies to link personal data with buying behaviour at the level of the individual customer. It works like this: you apply for the card to benefit from the organisation's promotion schemes; when you apply you give the organisation – a retailer or an airline, for example – your

Box 8.6

Examples: sources of internal digital data

- Sensor data: data from sensors embedded in products which allow them to be tracked from point of origin to consumer
- Scanner data: data from scanners such as electronic point of sale (EPOS) devices which scan machine readable codes on products and forms
- Interaction data: data interactions and/or communications, for example from cookies in websites, weblogs, information in email headers, and information on time, date, duration and location of phone calls
- Transaction data: data from in-person and online transactions.

personal details; each time you make a transaction and have your 'loyalty' card scanned, the personal details from the card are recorded and logged against that transaction and so against the purchases you made. Your individual record can be updated with these purchases. The organisation has a record of your actual buying behaviour (that is, not your claimed buying behaviour – which is the sort of data that is recorded in a survey) in your personal record in its database. The same process operates in online retail. You register to use the site and your activity on it is recorded and stored.

Your purchasing patterns can be analysed (for example, in a 'shopping basket analysis') and, on the basis of this past behaviour, the retailer can send you recommendations and alert you to offers on the types of items or brands that you buy. The retailer can also compare your purchase record with that of other customers and, where they detect similarities (say in the purchase of X), notify you, telling you that 'people who bought X also bought Y and Z'. 'Shopping basket analysis' can show what sets of products or brands are bought together, and which groups or segments of their customers buy which sets of products or brands, allowing them to target groups of customers with tailored offers, for example, and they can see from the database the uptake on these offers. Further, by examining trends in behaviour over time, the company can build models and algorithms to predict behaviour, sales volumes and revenue. This information can be used to understand, for example, how profitable different groups of customers or different types of outlet are, and what type of promotion works best for which group.

Data derived from loyalty cards, however, can be limited. While they give information about customer behaviour in the store, they do not give information on behaviour outside it (for which data from consumer panels may be useful); the demographic information provided may not in all cases be accurate; and people may hold more than one card for the same store (Passingham, 1998). Also, the customer may not use the card for every transaction. In addition, loyalty card data cannot be used to build a full picture of a store's or a vendor's customer base as some customers may refuse the offer of a card. It is a different matter when it comes to online buying activity, which does not rely on the loyalty card method although many sites request registration.

Data from customer interactions can be a source for secondary research, providing there is a lawful basis under data protection legislation for processing or using the data. The data can provide detailed current and historic information about *actual* (observed) rather than the *reported* buyer behaviour on which primary research such as a survey might report. In market research terms, the sort of questions these data are helpful in addressing are questions related to understanding customers, identifying emerging patterns and trends, planning and designing promotions and advertising campaigns, and developing new products, to name a few.

Industry Insight 8.5, although conducted many years ago, shows how the analysis of basic customer information recorded at the box offices of arts venues provided greater insight into customer behaviour than did customer surveys at the venues. It shows, too, how data can be shared and used for decision making by several parties (when there is permission to do so). We looked briefly at this example in Industry Insight 3.4 Ticketing talks; here we look in more detail.