How much is your company's data really worth?

Data pricing and valuation in the age of the data economy.

White Paper | September 2017

Contents

1.0 Introduction . . . **3**
1.1 Welcome to the data economy . . . 3
1.2 Executive summary . . . 4

2.0 Assigning value to data . . . **5**
2.1 Defining 'data' for the purpose of this paper . . . 5
2.2 Price, value and revenue opportunity . . . 6
2.3 Rational data valuation . . . 6

3.0 Factors that determine the value of data . . . **7**
3.1 Business value . . . 7
3.2 Incisiveness . . . 7
3.3 Uniqueness . . . 8
3.4 Company size . . . 8
3.5 Granularity . . . 9
3.6 Cost context . . . 9
3.7 Reliability . . . 10
3.8 Accepted usage and risk . . . 10
3.9 Freshness . . . 10
3.10 Commitment . . . 10
3.11 Collection cost . . . 10

4.0 Application to real world examples . . . **11**
4.1 Data typology . . . 11
4.2 Pricing aggregate data . . . 11
4.3 Pricing personalisation data . . . 12

5.0 Guidelines for commercialisation – Applying data valuation in market . . . **13**
5.1 Overview and purpose . . . 13
5.2 Data commercialisation principles . . . 13

6.0 Recommended actions – 6 Steps to getting started on Data Commercialisation . . . **15**

1.0 Introduction

1.1 WELCOME TO THE DATA ECONOMY

Earlier this year, The Economist proclaimed "the world's most valuable resource is longer oil, but data." The Economist was speaking about a new global economy – the data economy. An economy that is about collecting, storing and analysing streams of data that arise from every aspect of human activity.

Steve Millward,
Chief Analytics Officer, Data Republic
Sign-off from author

The most valuable companies in the world are so valued because of the data those companies control. While initially seen as using data for advertising, Google and Facebook have expanded the use of their user data to train voice, image recognition and AI software. Uber provides cheap taxi rides but also holds the biggest dataset of supply and demand data for personal transport.

This data is enabling these companies to better understand the world's economic, social, consumer and business infrastructures, as well as build data-driven decision-making into every aspect of their operation.

However, a crucial problem remains in that we don't have a unified, comprehensive way to value data to enable it to be freely traded.

All the things we can do with data–improved customer experiences through better targeting and personalisation, streamline businesses processes, or even find cures for disease or provide better hospital care–can't eventuate without a functioning data economy where data has clear, measurable value.

And because we can't value data, we're faced with three key problems:

- Businesses are sitting on massive, unvalued, under-utilised databases leading to unrecognised shareholder value.
- Revenue models based on the use of data aren't being implemented leading to missed opportunities.
- We're missing out on the extra demand and activity that attributing value to data could create as a data economy flourishes.

So why hasn't this happened yet? Why don't we have a standard for valuing data?

After spending more than twenty years working with data in and alongside big business, I think there are two main reasons.

The first is that most businesses, although they increasingly view data as an asset, see it mainly as a back-office issue. Something to be maintained, like a tool. However, this is starting to change as boards begin to recognise the untapped value.

The second issue is that businesses simply don't know how to value data. No general accounting principles exist to put data on the books and no marketplaces have existed to prove that valuation.

That second point is especially tricky. Data is unique. Its value differs for whatever entity accesses or maintains it and as a result, creating a valuation figure is hard. Even more unusual is the realisation that the value of data doesn't decline with use and is usually variable.

This is one economic basis for exchanging data on a marketplace. Unlike other assets like iron or coal that have 'intrinsic' or transparent value, there isn't a single set of unique, accounting standards or methods by which data can be valued from an economic perspective and traded. This obfuscates its value.

The first step in the journey towards bringing data onto the balance sheet is for businesses to evaluate and treat data as an asset, as they would their brand.

The aim of this paper is to describe how businesses can determine the market price of their data assets when used by other organisations and the resulting total opportunity for revenue derived from data commercialisation.

It is my firm belief that as more and more businesses embark on practical data valuation processes, the true potential of data in our economy and society will begin to be realised.

So, what are we waiting for? Let's get started.

1.2 EXECUTIVE SUMMARY

- Data possesses unusual characteristics that makes it a unique type of asset. The process by which we value it therefore requires novel approaches.
- The value of data can be segmented into two distinct classes: buyer value and seller value.
- There are two types of metadata that can be used to value data;
  - The factors that determine the value of data to create the inputs into a valuation model
  - Examples of data pricing in established markets that can be used to calibrate the parameters within the model
- There are eleven metadata factors that can be used to determine the value of data, broken out into business needs, or demand factors, and data quality, or supply factors
- Operationalising a data commercialisation strategy necessitates the development and management of new business practices to drive long term success

2.0 Assigning value to data

Data is now one of the most valuable assets an organisation can have. Not only has the consumer journey been digitised, so have business operations and business-to-business interactions. Each engagement or transaction, whether at a consumer or business level, generates data.

If we combine and analyse enough of these data points, we unlock powerful insights on everything from customer buying behaviour to our own business operations.

These insights can then inform performance indicators, investment decisions and market opportunity. Because more data points across the customer journey and business decision-making process can be captured and analysed, data becomes an intrinsic asset class for the organisation.

Although powerful, data is a particularly difficult asset to value.

There are several unusual characteristics that makes data a unique type of asset:

» The value of data is wholly determined by the user, meaning that both the perceived and measured value varies greatly between different users.

» Data inventory doesn't decline after being sold. Data can be licensed to another user and the original owner of the data still retains ownership. Data is often renewed on a minute-by-minute basis, meaning its value is refreshed regularly. The same data can be licensed to multiple users at a single time.

» Data is a raw material, and like other raw materials, it can be refined and improved by processing and generation. However, data currently has no open marketplace for trade, as is the case with iron ore, oil or gold, therefore the value of data can be obfuscated.

» Data is an illiquid asset. It is complex and time-consuming to make data transactions, resulting in a small proportion of the world's data being traded and assigned any tangible value.

2.1 DEFINING 'DATA' FOR THE PURPOSE OF THIS PAPER

Information has always been highly valued throughout human history. However, this paper is primarily concerned with the subset of information that fits within a modern description of 'data' as held by an organisation about the people it interacts with. In this paper, we see a person as the basic unit that data relates to. There are, of course, large volumes of other data unrelated to people but for now these are excluded.

It's easiest to think of data in relation to a person's activity. For most organisations collecting person data, this will be transactional data (CRM, sales), preference data (products, services and timings) and behavioural data (web, app, social analytics). The EFTPOS charge to your customer at the point of sale is the transactional data that will be recorded, whereas the customer's behaviour leading up to the sale (a Google search, a visit to the website, when they shopped at your store) is the behavioural and preference data.

Engaging with data on these terms is far more relevant and actionable than to simply call it "big data". We can of course summarise data as a "collection of facts about people" for brevity.

This is a loose description because:

» The size of the collection is hard to define as it has no lower limit. Generally a set of data covers a large number of instances of similar events or activities.
» The term "facts" is more directional than absolute. Many useful data may be indications only and are tested and proven by data scientists. The factuality of data will vary depending on the intended use.

Beyond this, we see some clear definitions in the broad types of data that companies now utilise.

Figure 1 – Types of data defined by ownerships and collection purpose

FIRST PARTY DATA	SECOND PARTY DATA	THIRD PARTY DATA
First party data is the information that is collected by an organisation directly from its customers as part of its usual business.	Second party data comes directly from another organisation. Second party data deals are made directly between the two parties and can include specific data points, audiences or segments. The data is often a by-product of the seller's usual business.	Data that is purchased pre-collected from an external provider or aggregator and can be compiled from a wide range of sources. The data is collected for the express purpose of commercialisation.

2.2 PRICE, VALUE AND REVENUE OPPORTUNITY

The value of data can be segmented into two distinct classes—the first being the buyer value.

Price is the dollar amount at which a vendor of data is willing to sell. If this price is too high and the data is not sold, then the price exceeds the buyer value of the data as perceived by the market.

The revenue opportunity is the total revenue of a data asset if sold rationally to the whole market at the optimum price. Due to inefficiencies in the market, realised revenue is usually less than the revenue opportunity

A second definition of value then relates to the seller value (Sv), which is the way that the seller attributes value for data sold within its business. At present this is the sum of all revenues (S R) minus the costs of selling (Sc).

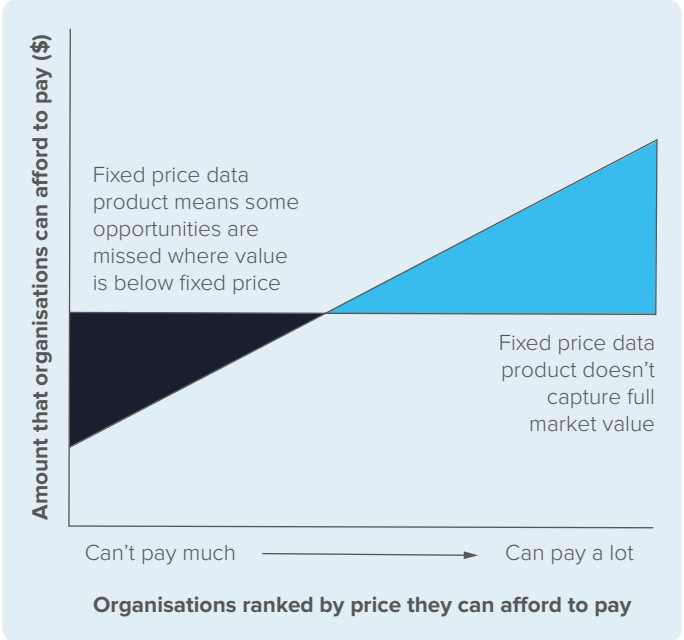
Sv = S R — Sc

Traditionally the cost of selling was high due to the entire process being managed internally. This cost is now falling rapidly due to the establishment of data marketplaces that facilitate inter-organisation data access, in turn unlocking new datasets.
Data will eventually be valued on balance sheets as an intangible asset that generates a level of revenue and will drive increased company valuations at prevailing price to earnings ratios.

2.3 RATIONAL DATA VALUATION

Historically, most businesses that sell data have aimed for value-based pricing where the price charged is as high as possible for each user.

Figure 1 – Fixed price for data reduces the total revenue from data



Value-based pricing is not unique to the data industry, with airlines and utilities as common examples of industries that set price based on demand and value.

Data businesses have typically attempted to prepare 'price guides' that determine the value of data for different users and for different purposes. Complications arise in this model due to a lack of transparency on the value of the data to each user. The variables that determine what a set of data is worth to each user, particularly with regards to second party data, is not widely available.

Each user has a variety of ways to measure the impact of data on their operations or investment decisions, and there are differing levels of willingness to share that information. This has traditionally made it difficult for data marketplaces to flourish because of the disconnection between seller-determined value and buyer-determined value.

To bring transparency and clarity to this topic, we can look at two types of metadata that can be used to value data:

- The factors that determine the value of data to create the inputs into a valuation model.
- Examples of data pricing in established markets that can be used to calibrate the parameters within the model.

By understanding the types of metadata useful to valuing data, you will be able to take these methods and start to apply it to your own organisation's data, enabling you to recognise the commercial opportunity.

3.0 Factors that determine the value of data

This section outlines eleven factors that can be used to determine the value of data. The factors have been split into two groups:

» Business need, or demand factors. These factors are determined by the user with the knowledge of their own business and the problem they are seeking to address.

» Data quality or supply factors. These factors determine the value of different types of data to address a user's business needs, and ultimately determine the effectiveness of the data. This can then be seen as its ultimate value.

The list below is in approximate order of importance, from highest to lowest.

Figure 3 – Dimensions affecting data price

BUSINESS NEED (DEMAND)	DATA QUALITY (SUPPLY)
Business Value	Uniqueness
Company Size	Incisiveness
Cost Context	Granularity
Commitment Level	Reliability
	Freshness
	Accepted Usage and Risk
	Collection Cost

3.1 BUSINESS VALUE

Business Value is the most important factor in determining value because it is arguably the only "fact" that is known in a data transaction. Business Value has two distinct limitations:

» The seller of data may have limited understanding of the value to the buyer's business leading to inaccurate pricing and lost opportunities (see section Rational data valuation).

» The user may also have difficulty in understanding the value that the data can bring to their business. For example:

- Data could be used across many different departments in the organisation but not all departments may be engaged in the valuation process. For example, data that can be used to personalise marketing communications could also be used by risk, fraud, support and pricing teams. In many instances, these teams are not as closely connected to each other as they should be.
- Realising the full value of data depends on how well it is applied or deployed and many companies have resource or technology limitations.
- Measurement of effectiveness is not always done to a high quality so the value of data may be underestimated.
- Establishing the business value of data is currently unclear for the buyer and seller and can lead to a protracted negotiation period, which leaves neither side feeling satisfied.

3.2 INCISIVENESS

Incisiveness means understanding how true the data is to the intent of the customer. As data exchanges have started to become more commonplace, some standard patterns have emerged around preferred types of data and how they are best applied in an organisation. This application can either be on its own or when combined with other datasets.

Figure 4 – Table showing evolution of data types used by business

	BEFORE	NOW	FUTURE (ALREADY BEGUN)
Source	Research	Third party	Second party
Data type	Expressed	Observed (indirect)	Observed (direct)
Analysis type	Modelled	Modelled	Pass through
Incisiveness	Low	Medium	High

For example, it was common practice up until five years ago for shopping centres to survey one thousand people for one week as they left the centre to discover who they are, where they lived and what they purchased. This data was expressed in that it relied on people telling the data collector what they did. Today, most shopping centre managers are using bank transaction data to answer the same questions on hundreds of thousands of shoppers throughout the year. This transaction data is much more incisive across a far greater scale of buying behaviours.

Similarly, companies have used geo-demographic segmentations for over forty years to understand the sorts of products that different customers or prospects are likely to buy. However, data is increasingly available from marketplaces and websites that provide data on not only the products that each person is interested in, but more importantly, when they are in market for a particular product. Again, the data used in personalisation and targeting is becoming more incisive.

SIDEBAR - THE DECLINE OF THE GENERIC DATASET

Generic datasets have long been used to deal with marketing challenges related to acquisition and market knowledge. Age, gender and originally 'social class' based on occupation were used to create target markets in the twentieth century. By 1980 these approaches had evolved, using census data and new computing capabilities to create more granular geo-demographic classification systems. The systems in this group include ACORN, Mosaic, Landscape and GeoTribes. They split populations into a limited number of groups (10 to 200 clusters) to help marketers understand socio-economic, cultural and demographic differences between people.

Generic segmentations mean if Person A buys a new car, they are more likely to buy a BMW than the average person. People like Person A have a higher propensity (perhaps capacity) to buy the car; but only 2% of individuals might be in the market for a car at any one time, meaning the segment potential insight is quite broad.

Specific suggestions mean Person A is looking at Audis, so as a marketing executive in BMW or Mercedes or Audi I want to persuade Person A to buy my marque. And I can spend $100 communicating with them because they are worth it. Car manufacturers that cater to price-sensitive buyers can now make the decision to not spend marketing budget targeting Person A.

Incisiveness is important because it creates enormous efficiencies for businesses. If data can identify the 1% of the market that a marketer need to focus on rather than the 10% (containing both the 1% and 9% that are false positives), it saves the marketer cost in their business activities. Not only is the data worth 10x from its incisiveness but it reduces all other costs too, such as media, call centres, channel investment, etc.

It is somewhat ironic that as incisiveness increases the volume of data usually decreases. Multi-variate models that use an array of inputs are being replaced by more incisive data that reflects real intent. This is a continuing process, as marketing and other customer department budgets are increasingly under scrutiny and organisations seek to measure outcomes arising from their expenditure.

3.3 UNIQUENESS

The uniqueness of data is a supply-side factor that enables data vendors to control price. If a dataset is totally unique in its ability to address a particular business need, the value will be as high as the use cases can bear. However, like other products, market forces generally address imbalances between usefulness and uniqueness.

The evolution of specific data markets

» Absence of data. The market then looks for solutions by creating or acquiring data. This often starts by conducting small amounts of research and extrapolating the result across the country. For example, research may be conducted into customer demographic, geographic and buying behaviour for a hardware store in order to determine new store location opportunities. The initial research may be limited to existing customer data but the results will be extrapolated across other locations where similar buying behaviour can be assumed.

» Single, premium datasets. Google and Facebook currently have almost unique datasets that provide enormous competitive advantage to themselves and their users. This has led to many other organisations looking for alternative solutions. In most cases the most effective means of challenging single, premium datasets is by bringing together several other datasets into a single view.

» Multiple datasets. The data vendors compete on quality, reach, and comprehensiveness. Ultimately this can lead to competition on price alone although it may take up to 20 years to reach that point. Some vendors will drop out of the market as the combination of falling prices and market shares leads to uneconomic provision.

As business adoption of data marketplaces increases, where buyers can interact with and source data from multiple providers, there are two important observations for companies that are considering commercialising their data:

» First movers do have an advantage. Whilst first mover advantage in many industries has been comprehensively disproven, data is a sufficiently pure product that allows first movers to reach an entrenched market position. Once a dataset has been expensively integrated into the user's systems, the cost of switching to another provider is too high to be justified.

» Uniqueness can be created by joining datasets. There is an almost infinite range of combinations of second party data that have yet to be created, and enabling multiple datasets to come together can create tremendous commercial opportunities. A datathon was recently held with Melbourne Business School using data from Westpac, Qantas, grocery basket data and Victorian Department of Health. Over 250 data scientists uncovered a wide range of relationships between commercial data and health outcomes, including vaccinations, avoidable hospital visits and heart disease using data that had never before been brought together.

3.4 COMPANY SIZE

The assertion that large companies should pay more than small companies for the same dataset is sometimes controversial. In general, buyers of data have accepted this but the relationship between company size and data price is not linear.

A retailer with turnover of $1 billion or 1 million customers will not pay ten times more for data than a retailer with turnover of $100 million or 100,000 customers. In fact, a company that has ten times as many customers may only pay three times as much for a given dataset.

3.5 GRANULARITY

Granularity is a supply side factor that refers to the level of data resolution. High granularity data has many benefits for users:
Figure 6 – Comparison of high versus low data granularity

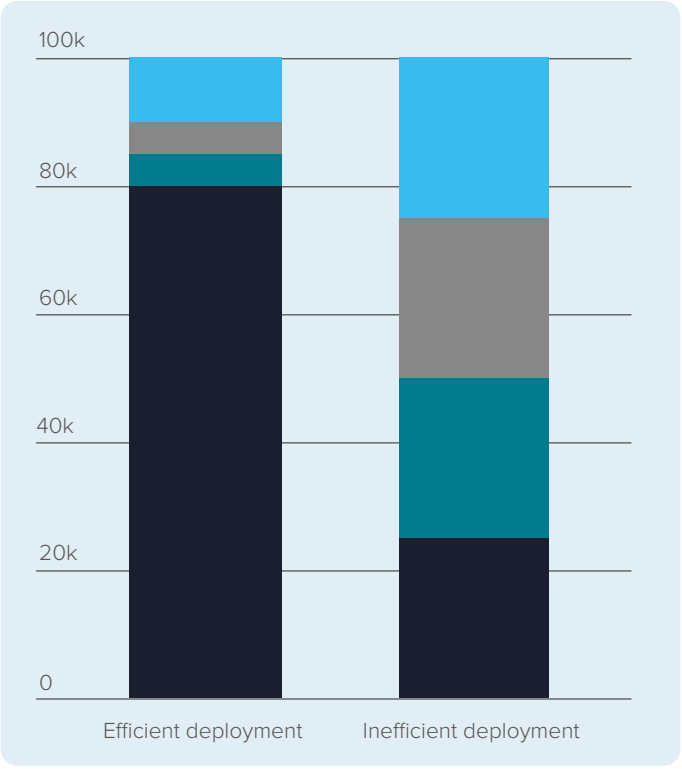
	LOW GRANULARITY	HIGH GRANULARITY
Insights	Customers spend $30 on average.	Female customers aged between 35 and 44 on the North Shore of Sydney spend $60, double the average.
Personalisation	If John bought a car, he would likely buy a luxury brand.	John is currently looking to buy an Audi A4 and has not arranged finance.
Modelling	Modelled outputs only give me a Yes/No likelihood.	Modelled outputs give me a score from 1 to 100.
Engagement decisions	I will target 50% of the market via social media.	I will target the top 6.4% of the market with an offer for a free gift.

Highly granular data is more valuable as it creates greater certainty and reduces wasted activity. However, there is a point at which higher granularity offers limited additional returns. For example, age or income bands are usually just as effective as actual age or income, assuming the bands are sufficiently narrow.

3.6 COST CONTEXT

Cost context refers to the additional cost that is incurred in applying data. Data is seldom deployed without additional cost. As noted in Business Value, the value of data is directly related to the value that a business applies to the solution it provides. Consider the diagram below.

Figure 7 – Diagram showing how data value relates to the overall costs of an application



A rational business will make a decision to buy data on the total cost of deploying the data against the benefit it brings. This sets a hard upper-limit on the value of the data. However, inefficient deployment of the data will reduce the amount that is available to pay for data.

In the diagram above the user can afford to pay $100,000 for both deployments, but in the inefficient deployment, the $100,000 is used up by cleaning, integration and interpretation, leaving $25,000 for data costs. Where the deployment is done efficiently, data can receive a greater share of the overall price.

Elements causing friction and increased cost of deployment can be helped by having specialist partners integrated into the data process. Two common causes of friction, namely cleaning and extract, transform, load (ETL), as well as distribution of results, can easily be addressed by integration with the right data technology providers.

Data science or the analysis component of using data is a relatively small proportion of the overall cost. Experienced analysts are usually able to generate outputs within a few days. Machine learning has the potential to advance this further, as insight gathering and learnings will be automatically applied to future datasets.

3.7 RELIABILITY

Reliability relates to how certain a user can be that the data is giving a true view of the subject. Whereas Granularity refers to how finely the data can be sliced, Reliability relates to quality. Questions that can help ascertain reliability are:

How truthful is the data? Expressed data from surveys always carries some margin of error.

» How well does this data cover the subject area? Is it national coverage or is it biased to some states leaving gaps in some areas? Is it skewed towards higher or lower income households? Is it skewed towards young people because it is collected via social media?

» How big is the sample? Is 10% of all customers enough? How much more would you pay for 20% or 40% coverage

» What is the lineage of the data? Where has it come from? Can the provenance of the data be proven? Anybody who has ever worked in marketing has received calls or emails from a third party trying to sell tens of thousands of prospect contact information. In most instances the data has come from another source or is out of date. Usually the uselessness of the data is only discovered after the sale

3.8 ACCEPTED USAGE AND RISK

The risk accepted by companies selling their data is usually determined by the uses they allow. The risk of moving data from one company to another is generally viewed as something of a binary gateway – if it cannot be done safely then it should not be done at all.

Advances are being made however in the technology department to allow companies to govern the permitted use of datasets and control the provisioning of data for value discovery. Data Republic's Senate Platform is one such tool which allows companies to securely exchange data with multiple entities under a single comprehensive legal framework and from one secure governance dashboard.

When the risks to data exchange are removed or minimised to such an extent that they are negligible, the end-use or application alone will be the driver of value.

3.9 FRESHNESS

Data freshness is a supply-side factor relating to how recently the data was collected. A company may wish to have intent data for its customers refreshed once per month and be successful in using that data to drive personalised offers. But when the company's competitor has weekly refreshes, it will be able to provide personalised offers before others.

As data collection, processing and storage technology has improved, refreshed data can now be made available in near real-time via Data Management Platform (DMP) solutions. Driven by the need for organisations to act quickly across multiple digital channels, DMPs have exploded over the past few years, lowering the technological barrier to entry and delivering cost-effective solutions for near real-time digital cookie data.

Freshness alone however may not stack up to other more significant factors in data valuation.

3.10 COMMITMENT

Most companies that sell data are willing to offer discounts for longer term contracts or company-wide deals. Longer term deals provide surety for both buyers and sellers.

Vendors of mission-critical data may be less willing to discount.

3.11 COLLECTION COST

Collection cost used to be a major driver of data value. Surveys had a fixed cost to complete and those costs were used to derive the final price.

A large proportion of data now used by companies, especially second party data, is collected as a by-product of another business process. Bank card transactional data is a by-product of offering banking services. Grocery basket data is a by-product of operating a supermarket. As noted in this paper, the value of data is now driven by many other factors and collection cost shouldn't be a consideration.

There are two important observations for vendors in relation to setting value based on cost:

» The cost of providing data should be looked at over the medium term as part of the lifetime value of a data product. Companies should make data easy to integrate and trial, and costs may exceed revenue in the short term.

» Data revenue is often added as a line item in business plans to get projects approved. However, the cost of monetising data is often very high, especially for companies that don't do this as their core business. Here is where opportunities exist for organisations to partner with other companies that do offer this as their core business.

4.0 Application to real world examples

4.1 DATA TYPOLOGY

There are two broad types of data that are useful to businesses:

» Aggregated data that provides insights on a market or group of people
» Personalised data that provides signals on an individual's intent or likelihood

These two types require distinct pricing models.

4.2 PRICING AGGREGATE DATA

Aggregate data are summaries of record-level data that provide a view of a market. Common examples include:

» A report from a market research survey that tells a brand about its customers.
» A summarised data cube that represents the aggregation of millions of rows of transaction data that can be used by a CFO to measure market share and sales uplift.
» A summary report from millions of rows of supermarket transaction data that can provide a brand manager with ideas for cross-promotion of their beer brand.

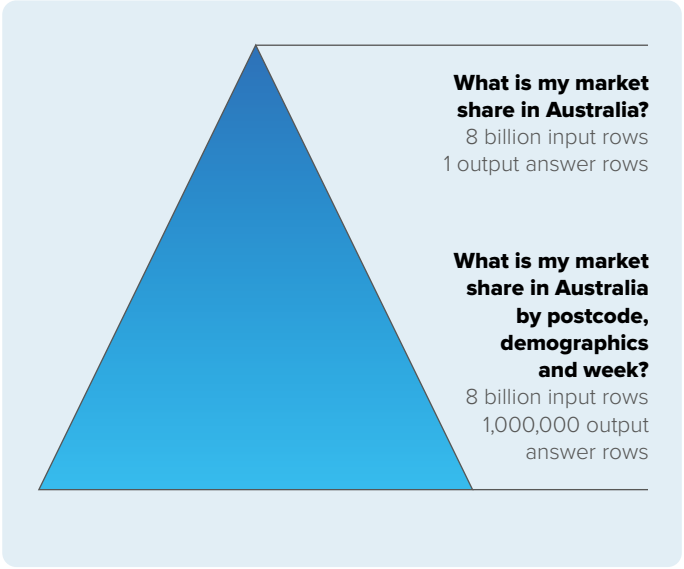
A key challenge in pricing aggregate data is whether to use inputs or outputs.

Pricing on inputs involves counting the number of individual rows that go into the analysis. This method is not ideal for pricing as it divorces the price from many of the factors outlined in Section 3, most notably Business Value, Incisiveness and Granularity.

Conversely, pricing on outputs provides closer alignment to Business Value. Buyers are able to engineer the outputs that they need to match a budget while still using the same underlying data.

Consider the figure below. Two questions have the same data input requirements but have vastly different output data.

Figure 8 – Comparison of using input versus output data size for pricing



For aggregate card transaction data, we have developed a points-based scoring system that allows the buyer to choose parameters against each dimension of data to meet the desired result. A version of this is available at datarepublic.com.

Figure 9 – Dimensions of pricing in aggregate card data

DIMENSIONS	CHOICES AND WEIGHTINGS
Location of card holder	Price is higher for more granular outputs, such as SA1 versus SA3 or Postcode.
Demographics of card holder	Price is higher for splitting outputs by gender and more detailed age bands, or by linking to other datasets such as geo-demographic segmentation.
Location of merchants	Price is higher for more detailed groups of merchant locations, such as SA1 or for defined commercial areas such as shopping centres.
Product category of merchants	Price is higher for fine categories such as "women's fast fashion" than for "fashion" as a whole.
Time	Price is higher for detailed analysis by hour or day part.
Date	Price is higher for more detailed analysis by day, for example to measure the impact on marketing share of a marketing campaign in a specific city.
Measures	Price may be adjusted for further filtering by number of transactions or transaction value, for example to enable RFM analysis.

The crucial outcome from any pricing model is to create a model that is based on science rather than opinion. A key strength of this approach is the ability of a buyer to engineer outputs based on budget and business need.

4.3 PRICING PERSONALISATION DATA

Personalisation data is best priced on a per record enriched basis. The challenge comes from setting the appropriate price per record. Over the last 30 years, data enrichment has typically been priced in the range of $20 to $800 per 1,000 records depending on the incisiveness of the data.

The factors outlined in Section 3 are shown below with an indication of the "multiplier effect" that it will have on the value of data. This is scored from 1 to 10, where an effect of 10 means that the value of data can rapidly escalate when this factor is very well matched. A multiplier close to 1 means that the factor generally has limited effect.

We typically start with a notional base value of 10c per record, or $100 CPM. This base value is then multiplied by each factor. This model has been in use by Data Republic for some time and it is important to note that these weightings will be revised as more metadata and more datasets become available.

Figure 10 – Dimensions and weightings

DIMENSIONS	MULTIPLIER	EXPLANATION
Business value	10x	The most important driver as it is directly linked to ability to pay.
Incisiveness	5x	Higher multiplier effect as it relates to how true the data is to the intent of the customer.
Uniqueness	5x	Important because it directly impacts how the data vendor can control pricing.
Granularity	3x	Highly granular data is more valuable as it creates greater certainty and reduces wasted activity.
Cost context	2x	This is lower for personalisation products as deployment is often a simpler process.
Reliability	2x	This is generally proven via trials that demonstrate business value. However, long-term assurance on availability of data is valuable.
Freshness	2x +	This is increasing in importance, as speed of execution becomes a point of difference.
Accepted usage and risk	1x	Weighted lower as there are many variables that influence its impact, and in the case of data exchanges, the risk can be removed.
Collection cost	1x	Not a useful factor in the open market.
Commitment	0.7x	Longer term commitments generally reduce the cost of data,

5.0 Guidelines for commercialisation – Applying data valuation in market

5.1 OVERVIEW AND PURPOSE

The purpose of this section is to provide guidance to companies interested in effectively valuing and commercialising data assets.

Having discussed, developed and implemented data commercialisation strategies with hundreds of companies who are intending to participate as contributors, partners or end users, Data Republic has accrued some unique perspectives and insights on what works and what doesn't when it comes to scalable data commercialisation.

In this section, we'll present a quick summary of our key findings.

5.2 DATA COMMERCIALISATION PRINCIPLES

Companies that are in the business of data commercialisation adopt specific practices to drive long-term success. In our experience at Data Republic, the following principles are crucial to effectively meet strategic objectives.

Principle 1 – Assess potential data products on 'lifetime value'

» For data businesses, the most valuable data products are those that are built once, sold to many different companies and licensed annually so that annuity streams are created.
» The licensing of data to create annuity streams is usually known as 'recurring revenue'. It is the foundation of creating a high-growth data business.
» The decision on pricing for any data product should be considered on the lifetime value of the product. Data products should be assessed on a net present value (NPV) basis.

Principle 2 – Assess new End Users on 'lifetime value'

» Acquisition of new clients is crucial but hard, especially when incumbents are present. Initial engagements may operate at a loss but the value created from a new client over their lifetime is significant.
  · Annual license fees mean that the NPV of a new client can be many times the initial value of the initial engagements.
  · Newly acquired clients are generally 'sticky' in that once they have chosen a data provider they will return for more data to tackle new projects or delve more deeply into previous analytics.
» Returns from acquiring new clients need to be examined on a multi-year timeframe, not just on the initial project.

Principle 3 – Manage the data product lifecycle

» Data products being developed by external parties require up-front investment by partner developers, so they all begin as loss-making activities. Data products only make profit once built, launched and sold to a number of clients.
» Data product developers will make the decision to invest in building a data product based on their up-front cost versus long-term return.
» Data marketplaces and Data Contributors should remove obstacles from the data development process to enable as many partners as possible to take products to market. As an example, Data Republic is creating tools and documentation to enable transactional data products to be built as quickly as possible.

Principle 4 – Low value products have an important place in the product mix

» Volume is as important as value. Many high revenue and profitable data businesses sell simple data products at low cost but high volume.
» Low cost data products are an important hook to acquire new clients. These clients then go on to buy more products and pay repeating annual licenses.

Principle 5 – Commercial opportunities should be seized today

» Data is a unique type of product in that its inventory doesn't diminish when sold.
» Therefore, revenue is maximised when opportunities are taken, even if the price of data is less than desired.
» Discounted sales often provide leverage to sell other data, get longer-term contracts or ask for published case studies. All of these create additional value.
» The risk of inflexible commercial terms is that Data Contributors are unable to convert enough opportunities and revenue targets are missed.

Principle 6 - Proofs of concept are necessary even if loss-making

» Proofs of concept (POCs) are necessary to drive long-term outcomes. Some POCs will not deliver results and will result in a net loss. However, overall POCs serve to reduce risk for the buyer:
  · The yield from data application is not always clear to data users. For example, they may already have models using other data, and are unsure of additional lift.
  · Corporate users of data are often unwilling to take risks in switching from an adequate supplier to another supplier that potentially offers a better solution. The economic and personal risk of making mistakes discourages switching, so a POC reduces that risk.
» A service model and fee structure for POC work is a key dimension for driving revenue at commercial data companies.

Principle 7 – Data Contributors become external client-servicing business

» Clients will need to come first within the risk and regulatory framework of Data Contributors.
» End-users, who are paying for data, will increasingly engage directly with Data Contributors, often online, and sometimes in person. They will expect levels of service that they receive from other service businesses, including timeliness of response, flexibility and teamwork.
» Clients also include partners, who are investing in data product development and are looking for long-term surety of opportunities. This may require longer contracts, flexibility on pricing and other forms of support.
» Consideration of team structure and team members is required to maximise opportunity.

Principle 8 – The sales and pricing process is complex

» Once mature, each Contributor's data commercialisation business will be well structured and predictable. In the meantime, flexibility is required to make client and financial decisions that carry commercial risk, even though those decisions are made within the overarching strategic objevctives.
» Data marketplaces should share a detailed data commercialisation plan with Data Contributors, including partners, target companies and revenue. This will enable both companies to agree on shared objectives and shared actions.
» As an example, Data Republic is aiming for list pricing for data in the medium term. However, in the short term larger clients will require negotiation over pricing, and this may be protracted. This will require skilled negotiation that can best be achieved by Data Republic and Data Contributors working together on strategy.

6.0 Recommended actions – 6 Steps to getting started on Data Commercialisation

ACTION 1
Set data commercialisation objectives

We advocate that Data Contributors should determine and socialise clear strategic objectives for commercialising data in the medium term to avoid conflicting decisions. These overarching objectives will provide the framework for tactical decisions on pricing, data products, staffing and client engagement.

» Set strategic objectives for the data commercialisation strategy as an overarching framework.
» Set team ethos to be that of a new market entrant with ambitious growth targets. Foster commercial drive to let teams flourish.
» Publish and share within the organisation so that stakeholders understand the objectives and tactics to meet objectives

Outlined below are example objectives for each stakeholder group in relation to commercialising data.

Data Contributor strategic objectives:

» Be the leading company in your sector, in Australia, as measured by performance versus competitors, by using its own and second party data to drive customer personalisation.
» Drive better outcomes from all analytical activities by using improved data sources.
» Meet financial targets for data commercialisation and become a profit-centre or recycle data revenue into new data-driven initiatives.

Examples of Data Contributor financial targets for data commercialisation

» Data commercialisation becomes a cost neutral service in the short-term (end of 2018).
» Data commercialisation becomes a profit centre allowing re-investment in use of data and cost offset in the medium-term ($Xm in gross data revenue by 2019/20, with net revenue share of $0.5Xm and profit of $0.4Xm).

ACTION 2
Set pricing or engagement frameworks and policies

» Set guideline pricing frameworks for different industries and use cases.
» Agree that these can be used externally as a starting point for discussion including ranges that can be indicated on data marketplace sites.
» Set policies for pricing that incorporate the framework.
» Be clear in the policies that the strategic objectives are the end goal and that there is flexibility within guidelines for the right deal.

ACTION 3
Set appropriate commercial decision-making frameworks

» Appoint a single person accountable for agreeing on variations in standard pricing, with backup people when the key accountable person is unavailable.
» Set a framework for agreeing on variations e.g. size of client. Data Republic has made a pricing model framework available on its website for a range of different datasets.
» Avoid committee thinking on variations. It should be a scientific process.

ACTION 4
Align team KPIs

» Align team KPIs and rewards with strategic objectives including revenue targets.
» Be clear that the team has flexibility and will be supported by senior management.
» Balance these KPIs against wider business objectives if accountable stakeholders have other responsibilities.
» Create shared goals to enable 'one team' thinking on commercialisation.
» Share and agree on prospect lists enabling assistance with forging opportunities, not just responding to late-stage interest from clients.

ACTION 5
Develop plan to address data quality

» Scope all possible options to address data quality.
» Write a detailed plan to explore each option and evaluate effectiveness.
» Document data quality measurement as a key business metric and make visible within the organisation.
» Assign resources based on client needs and opportunities.

ACTION 6
Review team requirements and structure

» Consider separation of functional roles across pricing, sales support, data operations, and observations of data protocols. Pricing and sales support in particular need commercial focus.
» Consider requirements for data commercialisation experience.
» Consider ring-fencing the commercialisation team with costs offset against revenue to create a profit centre. Align incentives with commercial outcomes.

THIS EBOOK IS INTENDED ONLY AS A BASIC OVERVIEW AND INTRODUCTION TO DATA VALUATION.

If you'd like to find out more, you can read about our platform at www.datarepublic.com

Alternatively, reach out to us directly – we are happy to answer any questions. Data Republic has offices in Australia and the United States and clients all over the world. You can contact us at:

Data Republic Sydney Office:
Level 1, 50 Bridge Street Sydney, NSW 2000 1300 138 254
enquiries@datarepublic.com

Data Republic San Francisco Office:
415 Stockton Street San Francisco, CA 94108 1-866-899-9886
enquiries@datarepublic.com

Non-profit organizations, charities, social welfare organizations and enterprises interested in the value of data for corporate social responsibility can learn more about The Minerva Collective – Data Republic's affiliated not-for-profit – at http://www.minervacollective.org/