Business Analytics and Statistics 1st edition Black Solutions Manual

SKU: 9780730363330-SOLUTIONS

Solutions manual

to accompany

Business analytics and statistics

1st edition

by

Black et al.

WILEY

Chapter 1: Data and business analytics

Review problems

1.1 What is a business strategy?

A business strategy is a long-term plan that sets out how a business will achieve its goals. Formulating a business strategy therefore first requires defining the business's goals. This requires an understanding of:

- the objectives of the owners and other stakeholders in the business (e.g. the community, creditors, regulators and interest groups)
- the internal and external resources that the business can influence or control
- the internal and external environment in which the business operates.

Once the goals are defined, formulating a plan to achieve them requires numerous decisions on questions and choices such as:

- the products and services the business will offer
- the market segments to target
- the value the product offers customers
- the features that differentiate the business's offerings from competitors.

Any business preparing a business strategy will face a list of such questions, some relatively broad and others more specific. Unfortunately, business decisions are rarely made with access to complete and perfect information. Rather, answering questions such as these requires information about many factors related to the business and the environment in which it operates. This information is often obtained by gathering, processing and analysing data.

1.2 Differentiate between data and information.

Many people use the words data and information interchangeably. In the context of business analytics and statistics, it is important to differentiate them. Data are simply quantitative or qualitative values. Qualitative data may be nominal or ordinal and quantitative data may be discrete or continuous. Data, and especially large data sets, are often of little practical value.

However, data can be gathered, processed, stored, manipulated, analysed and tested. These steps, in combination with organising, contextualising and presenting the outcomes of these steps, turn data into information. Provided the data have been gathered and processed using valid statistical methods, that information can then be used to inform business decisions.

1.3 What is the difference between business analytics and business intelligence?

Business analytics is an approach to decision making that is informed by analysis of quantitative data. This analysis may range from simple counting, rule checking and arithmetic through to data mining and the use of regression models to predict future behaviours. To make the outcomes of such analysis useful for decision making, the results are often presented as charts, tables, interactive dashboards and so on. These outputs are collectively known as *business intelligence* and are used to display, examine and explore data. Business analytics will often be performed to meet the needs of decision makers and those decision makers will access the information via business intelligence.

1.4 Give a specific example of data that might be gathered from each of the following business disciplines: finance, human resources, marketing, production and management. An example in the marketing area might be 'the number of sales per month by each salesperson'.

Example in Finance: data on Price/Earnings Ratio can be very useful to financial investors since it facilitates them in fair market evaluation of stocks.

Example in Human Resources: A HR or organisational manager might be interested to collect data on staff turnover as high employee's turnover ratio can reduce the company's efficiency.

Example in Marketing: Marketers may be interested in existing or potential customers' addresses and purchase history in order to promote a service or product.

Example in Production: a CEO or a manager of a manufacturing organisation is interested in information about the cost of production to ensure the product meets it cost target.

Example in Management: a manager might be interested in operational information (such as cost, revenue, efficiency, etc.) of the firm in order to establish strategic goals.

1.5 For each of the following companies, give examples of data that could be gathered and what purpose the data would serve: Bluescope Steel, AAMI, Jetstar, IKEA, Telstra, ANZ Bank, Sydney City Council, and Black and White Taxis.

Bluescope Steel: collecting data on international steel prices, current international competitors, and new potential industry clients so they can set their operational and production goals and strategies

AAMI: data on the insurance products prices and the type of products offered by AAMI's competitors to set price and marketing strategies.

Jetstar: Data on travelling passenger's characteristics and jobs so more tailored flights/routs can be arranged.

IKEA: data on wholesale suppliers that can offer cheaper products to IKEA so IKEA would be able to maintain low price strategy.

Telstra: monitoring data on complaints received by the call centre on daily basis to measure customer satisfaction.

ANZ bank: customer survey data on internet banking to improve their online banking services.

Sydney City Council: Data on infrastructure or transport to manage resources in optimum way.

Black and White Taxis: data on peak hours and week days to ensure enough shifts and taxis are made available to customers.

1.6 Give an example of *descriptive* statistics in the music recording industry. Give an example of how *inferential* statistics could be used in the recording industry. Do you think data mining has a role in the recording industry? What role?

Descriptive statistics in recorded music industry:

(1) ARIA total sales of CDs this week, number of artists under contract to a record company at a given time.

(2) Total dollars spent on advertising last month to promote an album.

(3) Number of units produced in a day.

(4) Number of retail outlets selling the company's products.

Inferential statistics in recorded music industry:

(1) Measurement of the amount spent per month on recorded music for a few consumers then use that figure to infer the amount spent for the population.

(2) Determination of market share for rap music by randomly selecting a sample of 500 purchasers of recorded music.

(3) Determination of top ten single records by sampling the number of requests at a few radio stations.

(4) Estimation of the average length of a single recording by taking a sample of records and measuring them.

The difference between descriptive and inferential statistics lies mainly in the coverage of the data. These descriptive examples all gather data from every item in the population about which the description is being made. For example, Sony BMG measures the sales on *all* its compact discs for a week and reports the total.

In each of the inferential statistics examples, a *sample* of the population is taken and the population value is estimated or inferred from the sample. For example, it may be practically impossible to determine the proportion of buyers who prefer rap music. However, a random sample of buyers can be contacted and interviewed for music preference. Population market share can be inferred from the results. The results can be inferred to population market share.

Data mining in the recording industry:

The recording industry has a long history of collecting and analysing data, perhaps most easily seen in the 'charts' that have for many decades reported the most popular songs of the week. The charts have long reflected both physical sales (determined by sampling selected music stores) and radio air play. The charts have now been adapted to also reflect online sales and streaming.

The recording industry can access data from traditional radio stations; online radio stations; subscription and free streaming services such as Spotify, iTunes and Youtube; music video channels; concert venues; physical and online 'record' stores; search engines and numerous other sources. In addition, data directly related to music consumption can be coupled with broader data collected about music fans. The recording industry can therefore explore, analyse and respond to both patterns/trends in the market and individual behaviours. Due to the enormous quantities of data generated, data mining approaches are appropriate to intelligently search for meaning in the data.

The data generated through these avenues already influences what songs are written, recorded and released. Whereas historically many music industry decisions were made by producers and 'A&R' people in recording companies, today business analytics help the recording industry decide how to market particular music to particular market segments or even directly to individuals. The data collected informs the 'charts' as well as the use of music in ads and TV shows and the recipients of industry awards.

The US online streaming service Pandora, for example, uses its 'Music Genome' to analyse hundreds of attributes of each song to identify songs with similar traits in an effort to tailor an individual's playlist to match their preferences.

Data mining could help the recording industry navigate and benefit from the changes to its traditional business model, including the opportunities and threats variously presented by subscription and ad-funded streaming, piracy, file sharing, independent music publishing, and other changes.

While 'talent' has long been said to be the music industry's most important asset, 'data' must now rank alongside it.

1.7 Suppose you are an operations manager for a plant that manufactures batteries. Give an example of how you could use *descriptive* statistics to make better managerial decisions. Give an example of how you could use *inferential* statistics to make better managerial decisions.

Descriptive statistics in recorded music industry:

(1) Total number of worker hours per plant per week — helps management understand labour costs, work allocation, productivity, etc.

(2) Company sales volume of batteries in a year — helps management decide if the product is profitable, how much to advertise in coming year, compare to costs to determine profitability.

(3) Total amount of sulphuric acid purchased per month for use in battery production — can be used by management to study wasted inventory, scrap, etc.

Inferential Statistics:

(1) Take a sample of batteries and test them to determine the average shelf-life — use the sample average to reach conclusions about all batteries of this type. Management can then make labelling and advertising claims. They can compare these figures to the shelf-life of competing batteries.

(2) Take a sample of battery consumers and determine how many batteries they purchase per year. Make inference for the entire population — management can use this information to estimate market potential and penetration.

(3) Interview a random sample of production workers to determine attitude towards company management — management can use the survey results to ascertain employee morale and to direct efforts towards creating a more positive working environment which, hopefully, results in greater productivity.

1.8 Classify each of the following as nominal, ordinal, discrete or continuous data.

a. The RBA interest rate
b. The return from government bonds
c. The customer satisfaction ranking in a survey of a telecommunications company
d. The ASX 200 index
e. The number of tourists arriving in Australia each month
f. The airline a tourist flies by into Australia
g. The time to serve a customer in a bank queue

a. Continuous
b. Continuous
c. Ordinal
d. Continuous
e. Discrete
f. Discrete
g. Continues

1.9 Classify each of the following as nominal, ordinal, discrete or continuous data.

a. The ranking of a company on BRW's top 1000 list
b. The number of tickets sold at a cinema on any given night
c. The identification number on a questionnaire
d. Per capita income
e. The trade balance in dollars
f. Socioeconomic class (low, middle, upper)
g. Profit/loss in dollars
h. A company's ABN Standard & Poor's credit ratings of countries based on the following scale:

Rating	Grade
Highest quality	AAA
High quality	AA
Upper medium quality	A
Medium quality	BBB
Somewhat speculative	BB
Low quality, speculative	B
Low grade, default possible	CCC
Low grade, partial recovery possible	CC
Default, recovery unlikely	C

a. Ordinal
b. Discrete, but treated as continuous for analysis
c. Nominal
d. Continuous
e. Discrete, but treated as continuous for analysis
f. Ordinal
g. Discrete, but treated as continuous for analysis
h. Nominal
i. Ordinal

1.10 Powerkontrol Australia designs and manufactures power distribution switchboards and control centres for hospitals, bridges, airports, tunnels, highways and water treatment plants. Powerkontrol's director of marketing wants to determine client satisfaction with its products and services. He developed a questionnaire that yields a satisfaction score between 10 and 50 for participant responses. A random sample of 35 of the company's 900 clients is asked to complete a satisfaction survey. The satisfaction scores for the 35 participants are averaged to produce a mean satisfaction score.

a. What is the population for this study?
b. What is the sample for this study?
c. What is the statistic for this study?
d. What would be a parameter for this study?

a. The population for this study is the 900 clients of Powerkontrol Australia.

b. The sample is the randomly chosen group of 35 clients.

c. The statistic is the average satisfaction score for the sample of 35 clients.

d. The parameter is the average satisfaction score for all 900 clients of Powerkontrol Australia in the population.

1.11 Cricket Australia wants to run a marketing campaign to increase attendance at test matches. You have been hired as a consultant to conduct a survey and prepare a report on your findings.

a. What variables do you consider affect a person's interest in cricket test matches?
b. Design a questionnaire of 10 to 15 questions that will enable you to decide which section of the population the marketing campaign should target.

a. The variables affecting a person's interest in cricket test matches include: playing teams, significance of test match, ticket price, etc.

b.
- What is your age?
- What is your gender?
- What is your employment status: fulltime, part-time, unemployed, student?
- On how many days in the last 12 months did you attend test cricket matches?
- On how many days do you plan to attend test matches in the next 12 months?
- What do you consider to be a reasonable price for a ticket?
- How much are you willing to spend on food and beverages at the stadium?
- What kind of additional entertainment would you like to see in future test matches?
- How would you travel to the match? Tick all that apply: car, bus, train, other.
- Did you experience any transport problems in the past when attending test matches? If so, please specify.
- Is there anything in particular that would attract you to future test matches?

1.12 Define: (a) secondary data, (b) primary data, and (c) big data.

(a) Secondary data are data that have already been collected for some purpose and thus are already available, whether internal to the organisation or from some external source.
(b) Primary data are data collected for a specific purpose. Primary data are collected as part of a research effort, usually designed to inform a particular business need or decision.
(c) Big data refers to the deep and broad collections of data that arise from the ongoing collection of data through organic distributed processes (e.g. credit card or search engine use).

1.13 Compile a list of every piece of data you have provided to business organisations over the past 24 hours (choose a shorter or longer time if necessary). Do you think the data are valuable to those organisations? Why/why not?

Answers will vary. An extract from an example response might include:

- I listened to seven tracks on Spotify while getting ready for university, potentially providing Spotify with details of the music I listened to, the device I used and the time I listened.
- I logged into my university email account using my phone, potentially providing my mobile provider with details of my location and the domain I visited to check my email, as well as providing my university with a record that I accessed my email account.
- I purchased petrol for my car on the way to the train station, providing my credit card provider with a record of the time, location and amount of the purchase. I obtained a discount using an offer on a docket from my last supermarket shop, potentially providing the petrol station and supermarket with data about the effectiveness of their loyalty program.
- At the train station, I swiped my pass to board the train, providing data to the train service and triggering a top-up for my pre-paid train pass, thus providing further data to my credit card provider.

1.14 What are the three V's of big data?

- *Volume* — data is generated, captured and stored from numerous available sources, quickly building data sets.
- *Velocity* — many everyday activities result in the production of data that is automatically stored in real time (e.g. an ATM withdrawal).
- *Variety* — automatic data capture from so many sources means that data sets are both broad and deep — they cover numerous different issues and in some cases provide great detail.

1.15 Define data mining.

Data mining is the use of machine learning to explore extensive data sets in order to identify information and patterns and predict behaviours in ways that are not feasible using traditional statistical approaches. Data mining can discover similarities or shared characteristics in data, identify groups, find and model relationships, classify data, detecting anomalies and outliers and present findings.

1.16 Define machine learning in the context of data mining.

Machine learning is the use of algorithms that learn directly from data and thus can explore data in an iterative way. When an analysis approach that uses machine learning discovers something of interest in data, it will automatically perform further analyses based on those findings, and then further analyses based on those findings and so on. In contrast to conventional statistical approaches that examine an entire data set with a single test, machine learning can choose and analyse individual records in relation to nearby records and run multiple, iterative analyses, thus identifying relationships that would be virtually impossible to find using conventional techniques.