ANL201
DATA VISUALISATION FOR BUSINESS

STUDY GUIDE (5CU)

Course Development Team

Head of Programme	:	Dr James Tan
Course Developer(s)	:	Mr Eric Tay (Revised by Ms Jess Tan Wei Chin)
Production	:	Educational Technology & Production Team

Educational Technology & Production
Singapore University of Social Sciences
463 Clementi Road
Singapore 599494

Release V1.6

CONTENTS

COURSE GUIDE

STUDY UNIT 1

OVERVIEW OF BUSINESS PERFORMANCE MEASUREMENT

STUDY UNIT 2

FOUNDATION OF THE SCIENCE OF DATA VISUALISATION

STUDY UNIT 3

FOUNDATION OF THE ART OF DATA VISUALISATION

STUDY UNIT 4

BASIC DATA VISUALISATION TECHNIQUES

STUDY UNIT 5

ADVANCED DATA VISUALISATION TECHNIQUES

STUDY UNIT 6

BUSINESS PERFORMANCE DASHBOARD

COURSE GUIDE

1.1. Welcome

(Access video via iStudyGuide)

Welcome to the course ANL201 Data Visualisation for Business, a 5 credit unit (CU) course.

This Study Guide will be your personal learning resource to take you through the course learning journey. The guide is divided into two main sections – the Course Guide and Study Units.

The Course Guide describes the structure for the entire course and provides you with an overview of the Study Units. It serves as a roadmap of the different learning components within the course. This Course Guide contains important information regarding the course learning outcomes, learning materials and resources, assessment breakdown and additional course information.

1.2. Course Description and Aims

This course equips students with knowledge and skills to use Data Visualisation in measuring and monitoring business performance. At the end of this course, students will be competent in using Data Visualization software to create Business Performance Dashboard.

Course Structure

This course is a 5-credit unit course presented over 6 weeks.

There are six Study Units in this course. The following provides an overview of each Study Unit.

Study Unit 1 – Overview of Business Performance Measurement

This unit introduces the concepts of Business Performance Measurement, Balanced Scorecard and Strategy Map. Further, this unit will discuss the internalisation of Business Performance Measurement in an Organisation.

Study Unit 2 – Foundation of the Science of Data Visualisation

This unit covers the concepts, benefits and application of Data Visualisation, four basic stages of Data Visualisation process, the Semiotics of Data Visualisation and Perceptual Processing Model used on Data Visualisation. In addition, this unit will discuss about what data can represent, two fundamental forms of data, data attributes, data quality measurement levels and metadata.

Study Unit 3 – Foundation of the Art of Data Visualisation

This unit introduces four components of Data Visualisation. This unit will also discuss about the applications of the four components in Data Visualisation.

Study Unit 4 – Basic Data Visualisation Techniques

This unit covers several techniques in building Data Visualisation for categorical and time series data. This unit will also cover how to apply the techniques using Data Visualisation software.

Study Unit 5 – Advanced Data Visualisation Techniques

This unit introduces several techniques in building Data Visualisation for ad-hoc analysis environment. This unit will also introduce how to apply the techniques using Data Visualisation software.

Study Unit 6 – Business Performance Dashboard

This unit covers the concepts of business performance dashboard, and will discuss techniques in building business performance dashboard.

1.3. Learning Outcomes

Knowledge & Understanding (Theory Component)

By the end of this course, you should be able to:

- Describe an overview of business performance measurement strategy.
- Evaluate the appropriateness of business performance measure against business strategy adopted.
- Discuss the design principles of performance reports and dashboards.

Key Skills (Practical Component)

By the end of this course, you should be able to:

- Use appropriate visualisation based on given data.
- Summarise data through the use of business performance reports with the aid of data visualisation software.
- Develop business performance dashboards using data visualisation software.

1.4. Learning Material

The following is a list of the required learning materials to complete this course.

Required Textbook(s)

- Yau, N. (2013). Data Points: Visualization That Means Somethin. John Wiley & Sons.
- Murray, D. G. (2016). Tableau Your Data. John Wiley & Sons.

Other recommended study material (Optional)

The following learning materials may be required to complete the learning activities:

1	Software	Microsoft Office software, and Data Visualisation software, like Tableau Desktop.

1.5. Assessment Overview

The overall assessment weighting for this course is as follows:

Assessment	Description	Weight Allocation
Assignment 1	Pre-Course Quiz 01	2%
	Pre-Class Quiz 01	2%
	Pre-Class Quiz 02	2%
Assignment 2	Tutor-Marked Assignment	18%
Assignment 3	Group-Based Assignment	20%
Participation	Participation during Seminar	6%
Examination	End-Of-Course Assignment	50%
TOTAL		**100%**

SUSS's assessment strategy consists of two components, **Overall Continuous Assessment** (OCAS) and **Overall Examinable Component** (OES) that make up the overall course assessment score. Both components will be equally weighted: 50% OCAS and 50% OES.

(a) OCAS: Assignment 1 consists of one pre-course online quiz and two pre-class online quizzes weighted at 6%. Assignment 2 is Tutor-Marked Assignment weighted at 18%. Assignment 3 is Group-Based Assignment weighted at 20%. Participation is weighted at 6%. Both assignments and Participation will constitute 100% of OCAS.

(b) OES: The End-Of-Course Assignment is 100% of this component.

To be sure of a pass result you need to achieve scores of 40% in each component. Your overall rank score is the weighted average of both components.

Passing Mark:

To successfully pass the course, you must obtain a minimum passing mark of 40 percent for each of the two OCAS components. That is, students must obtain at least a mark of 40 percent for the combined assessments and also at least a mark of 40 percent for the OES Component. For detailed information on the Course grading policy, please refer to The Student Handbook ('Award of Grades' section under

Assessment and Examination Regulations). The Student Handbook is available from the Student Portal.

Non-graded Learning Activities:

Activities for the purpose of self-learning are present in each study unit. These learning activities are meant to enable you to assess your understanding and achievement of the learning outcomes. The type of activities can be in the form of Quiz, Review Questions, Application-Based Questions or similar. You are expected to complete the suggested activities either independently and/or in groups.

1.6. Course Schedule

To help monitor your study progress, you should pay special attention to your Course Schedule. It contains study unit related activities including Assignments, Self-assessments, and Examinations. Please refer to the Course Timetable in the Student Portal for the updated Course Schedule.

Note: You should always make it a point to check the Student Portal for any announcements and latest updates.

1.7. Learning Mode

The learning process for this course is structured along the following lines of learning:

(a) Self-study guided by the study guide units. Independent study will require at least 3 hours per week.

(b) Working on assignments, either individually or in groups.

(c) Classroom Seminar sessions (3 hours each session, 6 sessions in total).

iStudyGuide

You may be viewing the iStudyGuide version, which is the mobile version of the Study Guide. The iStudyGuide is developed to enhance your learning experience with interactive learning activities and engaging multimedia. Depending on the reader you are using to view the iStudyGuide, you will be able to personalise your learning with digital bookmarks, note-taking and highlight sections of the guide.

Interaction with Instructor and Fellow Students

Although flexible learning – learning at your own pace, space and time – is a hallmark at SUSS, you are encouraged to engage your instructor and fellow students in online discussion forums. Sharing of ideas through meaningful debates will help broaden your learning and crystallise your thinking.

Academic Integrity

As a student of SUSS, it is expected that you adhere to the academic standards stipulated in The Student Handbook, which contains important information regarding academic policies, academic integrity and course administration. It is necessary that you read and understand the information stipulated in the Student Handbook, prior to embarking on the course.

STUDY UNIT 1
OVERVIEW OF BUSINESS PERFORMANCE MEASUREMENT

Learning Outcomes

At the end of this unit, you are expected to:

- describe the concept of Business Performance Measurement
- explain the benefits of Business Performance Measurement
- explain what effective Business Performance Measurement is
- describe the principles of strategy-focused organisations
- explain the concepts of the Balanced Scorecard model
- explain the four Balanced Scorecard Perspectives
- explain the relationship between mission statement, vision statement and strategy within an organisation
- explain why strategy needs to be translated into strategic goals within the framework of a Balanced Scorecard model
- describe the Strategy Map and its cause-and-effect relationships
- construct a Strategy Map
- explain how Business Performance Measures are selected
- illustrate methodology in which organisation strategy is linked to Business Performance Measures
- construct appropriate Business Performance Measures
- set up the right targets for Business Performance Measures
- explain the success factors for implementing and sustaining Business Performance Measurement systems

Overview

This unit introduces the meaning, purpose and foundation of performance management.

1. Business Performance Measurement Concepts

WATCH

Business Performance Measurement Concepts
(Access video via iStudyGuide)

Business Performance Measurement is the process whereby an organisation establishes parameters within which resources, programmes, investments and acquisitions are to achieve the desired business results. Through this process, an organisation establishes criteria to determine quality of their activities, efficiency, and effectiveness in conducting their business operations.

1.1. What is Business Performance Measurement?

To *perform* means to initiate and execute a set of actions or activities. The translation of these actions into actual result, outcome or accomplishment is called *performance*. Hence, performance related actions are the means, whereas accomplishment is the result.

Measurement means determining the size and the amount or the degree of the item we are measuring. Performance measurement involves measuring or ascertaining either the business performance or the means that affect such outcome. Therefore, *Business Performance Measurement* is a process that involves collection, synthesis, delivery and graphical display of information related to *Business Performance Measures*, such as return of investment, customer satisfaction, etc.

The ability to understand, manage and improve business performance is largely dependent on the ability to successfully measure business performance. By measuring their business performance, organisations will be able to derive meaningful information from such measures and present them in report or graphical presentation for review or evaluation.

Business Performance Measurement is simple in definition but complex in implementation. The term seems to suggest that it is the use of measures to quantify business performance, but it is the establishment of the appropriate measures where complexity rears itself. In the Human Resource business function, organisations would like to measure efficiency and effectiveness of employees. In the Marketing business function, organisations would like to measure marketing effectiveness in terms of customer acquisitions and market penetrations. In the Business Operations function, organisations would like to measure business process efficiency, product and service quality, supplier efficiency and operation yields. Finally, in the Finance and Accounting business functions, organisations would like to measure business profitability and cash flow adequacy. Therefore, by nature, Business Performance Measurement covers multi-dimensional aspects in various industries.

Measuring performance in an organisation is a process and by itself provides limited use to an organisation. There is a need to understand current business performance level, plan future business performance targets, effect necessary changes and transformations within the organisation. This will help an organisation to achieve future business performance goals. This makes measuring business performance a strategic and critical task of an organisation. The ability to correctly measure relevant business performance enables an organisation to achieve the following benefits:

- Understand its inherent strengths
- Understand its inherent weakness
- Focus on customers' real needs
- Provide relevant training and upgrading skills to employees
- Understand its past and current financial health
- Predict future shortcomings that may arise if not addressed
- Improve and control critical processes
- Enhance product and service quality
- Motivate correct behaviours
- Align improvement activities to strategies

However, we must understand that only when relevant business performances are measured will the organisation reap the maximum benefit from the exercise.

Another important contribution of Business Performance Measurement is that it serves as both a driver and a feedback mechanism for an organisation to execute its strategies, fine-tune its strategic programmes, and gauge the level of achievement of the strategies. By measuring business performance that helps the organisation to fulfil its strategies, it also enhances the relevance of the Business Performance Measures and sustains business performance.

In acquiring such performance-based understanding, the interpretation of Business Performance Goals are essential. It is important to know what is considered normal or expected or abnormal. In other words, organisations need to improve their ability to adequately explain both the expected and the unexpected.

A set of basic understanding that provides the scope of what constitutes a Business Performance Measurement system is defined below:

Understanding 1

A holistic business performance information which is obtained by having measures in all three areas: organisation fundamentals, operating processes, and objectives.

Understanding 2

Business Performance Measurement provides information on what has happened and the projected future state of events.

Understanding 3

Business performance is often expressed in the form of measurable indicators. Such indicators provide either information on the state of events undertaken by the organisation, or information on the objectives and results achieved.

Understanding 4

Business performance is usually relative. Correct interpretation is often based on a carefully chosen reference or benchmark.

Understanding 5

Business Performance Measures are used to judge business performance that have to be regularly validated and updated, whenever required.

Understanding 6

Business performance and strategy should be closely linked. Strategy planning generates targets and objectives, which in turn requires execution of planned actions. The collective performance of all planned actions is an indication of the level of success an organisation has achieved in meeting its strategic intents.

Understanding 7

Business Performance Measurement is a means and not an end. Information provided by such measures should be used to derive the organisation's actions and continual improvements.

A general model of Business Performance Measurement is shown in Figure 1.1:

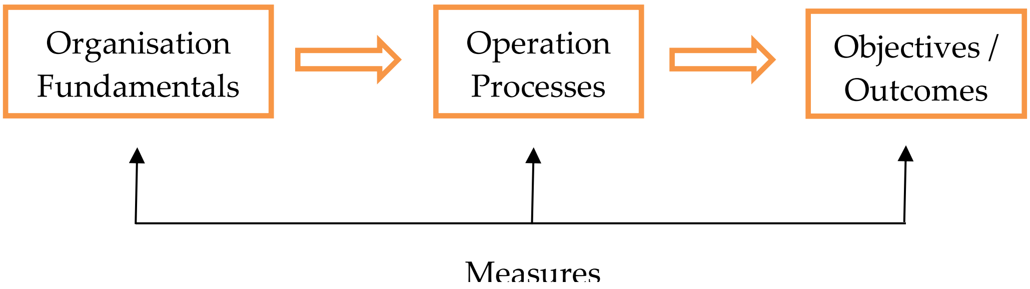


Figure 1.1: Business Performance Measurement Model

The aim of the above model is to link measures in three areas together so that the measures will give a holistic picture of the organisation's business performance in the past, the present, and the future. From the above diagram, we can see that organisation's objectives are dependent on how well the business operation processes are. Very often, how well such processes operate, depends a lot on the organisation's fundamentals, such as employee skills, productivity and motivation. This relationship leads to the need to have both *lagging* and *leading indicators* as measures. A lagging indicator provides information on what has happened, while as leading indicator provides information on projected future state of events.

An example of a lagging indicator is the Return on Investment (ROI). The ROI indicates the ratio of monetary gains an organisation has obtained in relation to the amount the organisation has invested. A good ROI value implies that the organisation has done well in the past year, but it does not provide information on how well the organisation is going to do in the following year. An example of leading indicator is customer satisfaction rating. A low satisfaction rate indicates a good possibility of increasing customer defections, thus a decline in sales and profits in the future.

A robust model of Business Performance Measurement requires measures to be established in all three areas so as to provide a holistic picture of organisation business performance now and the future. More discussion about such measures will be covered in Section 2.1.3 of this unit.

1.2. The Effectiveness of Business Performance Measurement

The following pointers are important considerations in effectively carrying out the Business Performance Measurement where measures are relevant:

- Do not confuse business objectives with business performance. Good objectives are the result of organisation exhibiting high business performance.
- Relevant measures are those that measure activities that can impact strategic objectives.
- Business processes are often made up of activities that involve many different business functions or departments. Therefore, measuring business process performance reflects more on the organisation's business performance than individual activities within a business function or department.
- Measurements should relate to customer's perspective of good business performance rather than to internal manager's perspective.
- The measuring method is as important as the selection of measures. Furthermore, measuring method is subjective and complex in nature.
- Measures should be used as a basis for improving business processes rather than finding faults.
- Measures shape employee behaviours, and therein lies the danger that employees seeking to improve a particular measure in a specific area which may have adverse consequences in other areas.
- Define measures precisely and execute measurements accurately within the context of justifiable costs and feasibility.

"If you can't measure it, you can't manage it"
Peter Drucker

2. Balanced Scorecard

WATCH

Balanced Scorecard
(Access video via iStudyGuide)

A study of 275 portfolio managers reported that the ability to execute a strategy was more important than the quality of the strategy itself. These managers cited strategy implementation as the most important factor in shaping management and corporate valuations. For the past three decades management theorists, consultants and the business press had been focusing on devising strategies that would generate superior business performance, and strategy formulation has never been more important.

2.1. Strategy-focused Organisation

The cover story on the 1999 issue of the Fortune magazine covered prominent CEO failures, concluded that the emphasis placed on strategy and vision created a fallacy that the right strategy was all it took to succeed. "In the majority of cases – we estimate 70 percent – the real problem isn't [bad strategy but] … bad execution," asserted the writers. Thus, with failure rates reported within the 70 to 90 percent bracket, it proves that strategy execution is more important than a good vision.

Below are reasons organisations face when implementing well-formulated strategies:

- Strategies – a unique and sustainable way that creates value – are changing but the system for measuring strategies have not kept pace.

- In an economy dominated by tangible assets, financial measurements were adequate to record investments in inventory, property, plant and equipment on a company's balance sheets. In today's economy, where intangible assets have become major sources of competitive advantage, organisations are encountering difficulties in managing what they cannot describe or measure.

- Many organisations operate under central control, through large functional departments. Strategy can be developed at the top and implemented in a centralised command-and-control culture. Change is incremental, so using slow reacting and tactical management control systems, such as budgeting, are inadequate for today's dynamic, rapidly changing business environment.

2.2. Balanced Scorecard

The development of an accounting system is synonymous with the industrial age where the only means of looking at performance was by looking at the financial results of organisations. For many years, business performance was judged based on financial results, usually at the end of one financial year. By reviewing business performance based on financial results, it is simply looking at events that are already over. This approach is considered as reactive instead of proactive.

In 1998, David Norton and Robert Kaplan (Harvard Business School) described how many world-class organisations have adopted Business Performance Measurement systems way beyond looking only at financial objectives alone. Kaplan and Norton correctly pointed out that measuring financial results alone was flawed in that they were documented after the event. Kaplan and Norton described it "like driving a car by looking at the white line in the rear view mirror."

In their quest for success, the above world-class organisations placed significant emphasis on non-financial measures. To provide a more holistic view of an organisation's overall performance in meeting the competition, and strategic objectives, Kaplan and Norton sought a better model of performance measurement. They came up with the concept of a balanced scorecard through their research.

A new approach to the strategic management was developed in the early 1990's by Kaplan and Norton. They named this system "The Balanced Scorecard" (BSC). The business performance measurement framework recognises some of the weaknesses and vagueness of previous management approaches. The BSC approach provides a clear prescription as to what an organisation should measure in order to balance the financial perspective.

Without a balanced scorecard, a business tends to be judged only by short-term financial results, and this may hide serious problems. For example, the reduction of short-term costs by deferring all maintenance expenditure. If the BSC was to include a question that probes the "percentage of maintenance completed on schedule," this safety issue would surface.

The BSC enables organisations to clarify their visions and strategies, and puts them into action. It provides feedback around both the internal business process and external objectives in order to continuously improve on strategic performance and results. When fully deployed, the BSC transforms strategic planning from an academic exercise into the nerve centre of enterprise.

Kaplan and Norton describe the innovation of BSC as such:

"The Balanced Scorecard retains traditional financial measures. But financial measures tell the story of past events, an adequate story for industrial age companies for which investments in long-term capabilities and customer relationships were not critical for success. These financial measures are inadequate, however, for guiding and evaluating the journey that information age companies must make to create future value through investment in customers, suppliers, employees, processes, technology and innovation."

The BSC suggests that we view an organisation from four perspectives or dimensions, and we develop measures, collect data and analyse the data relative to each of the following perspectives:

- The Learning and Growth Perspective, which is sometimes referred to as the Innovation and Development Perspective
- The Internal Process Perspective
- The Customer Perspective
- The Financial Perspective

2.3. The Four Balanced Scorecard Perspectives

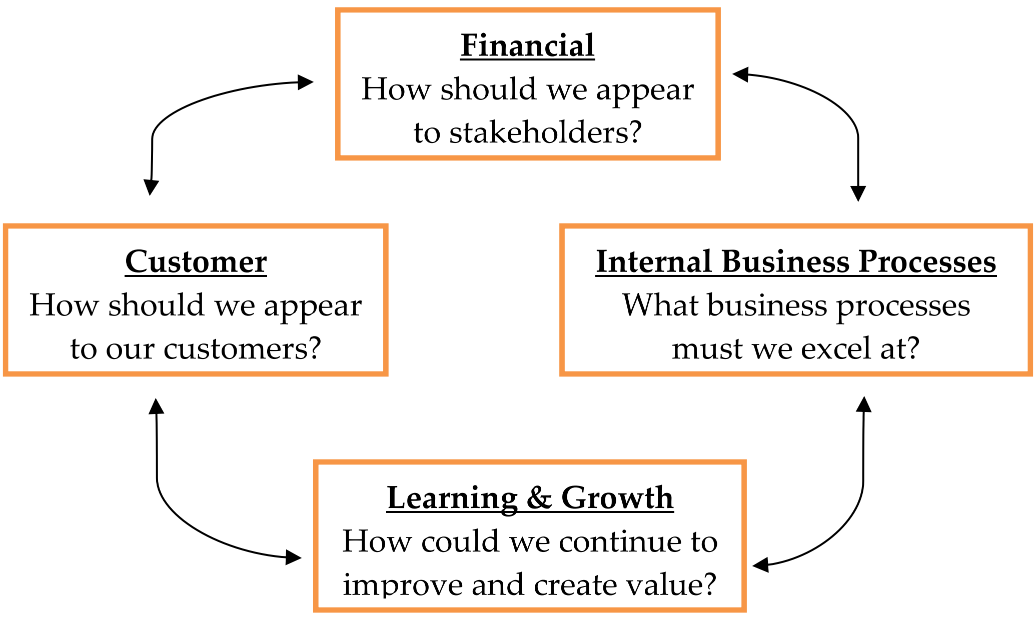


Figure 1.2: The Four Perspectives of the Balanced Scorecard

The BSC advocates a set of measures that relate to the four major perspectives, which are Financial, Customer, Internal Processes, and Learning & Growth, to determine the overall performance of an organisation. Customer, Internal Processes, and Learning

& Growth perspectives can be considered as operational measures that ultimately drive financial performance. The management will then have a complete view of an organisation's performance when viewing the overall performance in the current period compared with the previous period.

2.3.1. The Customer Perspective

Management philosophy in recent times has shown an increase in realisation of the importance of being customer focused (i.e. customer satisfaction) in any business. This is a leading indicator because if customers are not satisfied, they will eventually find other suppliers that will meet their needs. Poor performance from this perspective is thus a leading indicator of future decline in organisation's performance, even though the current financial outlook may look good.

Under the Customer Perspective, an organisation needs to determine what the customers require and think are important in terms of the products and services that the organisation provides them. Although customers' needs are varied and different from customer to customer, the needs also depend on the types of product and services offered. There are some basic customer requirements that are usually common across industries. These basic requirements relate to time, quality, service, functional performance, and value. Organisations have to establish goals for each key requirement. Measures must then be selected to gauge the organisation's success in achieving these goals.

For example, a goal related to quality might be to reduce customer returns in the coming year. A direct measure for this goal would be the percentage of customer returns. In developing measures for customer satisfaction, customers should be analysed in terms of the kind of customers they are, and the kind of processes for which we are providing a product or service to these customer groups.

2.3.2. The Internal Business Processes Perspective

Measures based upon internal process perspective allow managers to know how well their business is running, and whether its products and services meet customer requirements. Under this perspective, the organisation must identify key business processes that need to be excelled in terms of organisation's operations. Measures must then be established to monitor the performance of these key business processes.

For example, if one of the key business processes is new product introduction, then a possible measure can be the new product lead time, which is the time between the conception of new product idea and the actual introduction of the new product in the market.

As every organisation is unique in its composition, the measures for internal process perspective have to be carefully designed by those who know these processes most intimately.

2.3.3. The Learning and Growth Perspective

This perspective includes employee training and corporate cultural attitudes related to both the individual and corporate self-improvement. In a knowledge worker organisation, people – the only repository of knowledge – are the main resource. In the current climate of rapid technological change, it is becoming necessary for knowledge workers to be in a continuous learning mode.

Innovation usually refers to an organisation's ability to introduce new products and services, while *developmen*t usually refers to an organisation's ability to continually make improvements to its key processes and improve productivity. The organisation will then have to establish goals that relate to innovation, growth and development. Measures must be selected so that they can gauge the organisation's success in achieving the goals. For example, an organisation might set the following goal related to innovation -- to increase the number of patents filed per year. A measure for this goal might be the number of patents filed annually per research and development (R&D) employee.

Kaplan and Norton emphasised that learning is more than just training. It includes people like mentors and tutors within the organisation, as well as that ease of communication amongst workers that allows them to readily receive help when needed. Learning also includes acquisition of technological tools or high performance work systems.

At times, an organisation may find it difficult to hire new talents, at the same time there is a decline in training of existing employees. This is a leading indicator of "brain drain" that must be reversed. Measures can be put in place to guide managers in focusing training funds where they can help the most. Learning and growth constitute to the essential foundation for the success of any knowledge-worker organisation.

2.3.4. The Financial Perspective

Kaplan and Norton do not disregard the traditional need for financial data. Timely and accurate funding data will always be a priority, and managers will do whatever necessary to provide such data. In fact, often, there is more than enough handling and processing of financial data. The important point to take note of is that over emphasis on financials information can lead to an "unbalanced" view of performance.

Emphasising on non-financial related information such as risk assessment and cost-benefit data are equally important.

The measures for the other three perspectives that relate to customer, internal business process and people must be appropriately selected in order (for financial performance) to yield results. If the measures in the financial perspective are not showing gains while measures in the other three perspectives are, it could mean that the organisation has set the wrong goals.

For any profit-oriented organisation, the bottom-line profitability is important. Although the concept of the BSC is to take performance in many non-financial areas into consideration, the measures in the financial perspective must eventually show gains in order for the organisation to conclude that its overall performance has indeed improved. Typical financial objectives are growth, profitability and shareholder value. For instance, take growth as a goal, an organisation may set a goal that relates to sales growth. A measure for sales growth could then be the increase in sales dollar.

The BSC provides substantial focus, motivation and accountability in both profit and non-profit organisations. It provides the rationale for their existence (serving customers), and communicates to all stakeholders the outcome and performance drivers by which the organisation achieves its mission and strategic objectives.

3. Strategy Map

WATCH

Strategy Map
(Access video via iStudyGuide)

3.1. Vision, Mission and Strategy

Business performance measurement in itself serves limited purpose unless it provides linkages between strategy, strategic programmes, continuous improvements, and employee compensation in order for it to be considered an invaluable asset to the organisation. The term, "strategic business performance measurement," can be used to describe such a system.

The *Strategic Business Performance Measurement* is the process of executing strategies through planned activities that are intended to achieve certain strategic objectives. Measurable business performance in the form of critical performance indicators are then established through achieving such objectives to track advancement towards strategic objectives, and to drive necessary improvement actions that will bring the organisation towards its intended strategic objectives. A model of strategic business performance management is shown in Figure 1.3.

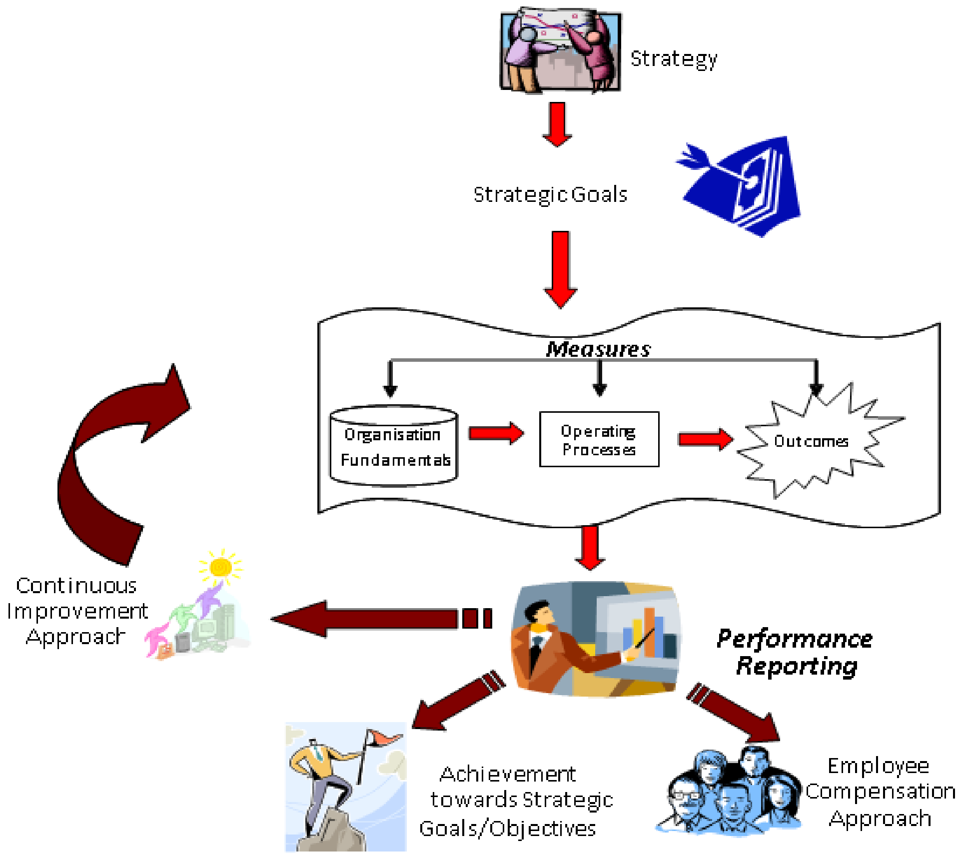


Figure 1.3: Strategic Business Performance Management Model

Strategic business performance measurement begins with an organisation establishing long-term strategy (by the CEO) to compete in the market. While strategy is a set of high-level plans, dictating how an organisation should operate and compete in the next few years, strategic objectives provide the direction necessary for employees to set their plans. This is the execution of strategy.

Once strategic objectives are established, business performance measures are then selected. That will provide good indication of the level of performance in key areas that have critical impact on achieving the strategic objectives. Once measures are in place, the business performance measurement system can then be deployed. Data will also be collected and collated. However, to provide a concise snapshot of the state of business performance to managers and key personnel, a focused business performance reporting system must be in place. Such systems are commonly called *business performance dashboards*.

An organisation carries out strategic planning in order to define its strategy and direction. Such process is often led by the CEO of the organisation. The pursuit for an organisational strategy is aimed at achieving its mission and vision in a medium-to long-term as shown in Figure 1.4.

Figure 1.4: Vision and Mission Statements

Although such statements provide a sense of direction, many people often could not draw a distinction between the mission statement and the vision statement. By definition, these statements can be defined in simple terms:

- A *Vision Statement* sets out the desired future state of the organisation.

- A *Mission Statement* is the reason for the organisation to exist.

Further understanding can be achieved by looking at the period which these accomplishments are made.

Strategy formulation and strategic planning are important tasks carried out by the management to steer the organisation to gain higher performance and greater profitability. Strategies formulated by organisations are usually relevant for three to five years, depending on the competitive landscape and technological changes.

A strategy is not formulated in a vacuum. It is usually guided by the organisation's mission, vision, and core values. Although the form of strategy being established is a product of competition and operating environment, the ideals of the strategy always revolve around the mission, the vision and the core values of an organisation.

While strategy planning is a key task, of equal importance is its implementation, many organisations fail not because they do not have good strategies, but because they do not execute their strategies effectively. A business performance measurement framework serves as an important role in supporting the execution of strategy within an organisation. Once a strategy is formulated by the CEO, it must be translated into strategic objectives and business performance measures. With the establishment of business performance measures, performance of various business activities within the organisation can be monitored towards achieving its strategic objectives.

Here are some examples of strategy established by different organisations:

- Increase growth of earnings to 15 percent in three years, and achieve a return on equity of 20 percent, so that the organisation will become the leader in its new products and attain lower costs than its competitors.

- Growth (through diversification) by merging with or acquiring firms in different, but related industries.

- Satisfy customers by providing quality products, reducing time taken to offer new products in the market, improving efficiency of all plants and processes, and building teamwork amongst employees, suppliers and dealers.

- Become the market leader by offering best-of-class products to customers.

- Protect and improve the firm's position as the number one brand globally. Build a strong momentum in growing the female consumer market. Continue to drive for increased margins through efficient inventory management.

- Global product strategy: same product and same marketing approach everywhere.

- Return to profitability: stabilise operation by downsizing, retrenchment and revitalising organisational resources and capabilities.

A few observations can be made by reviewing the strategies above. Most of the strategies have elements of improving the firm's financial performance, competitiveness and market position. To the top management, the broad interpretation of the strategies is quite obvious. However, to the lower management level, this understanding may be unclear or lacking.

3.2. Strategic Themes

Strategic Themes or *Strategic Thrusts* are the main high-level business strategies, which an organisation must excel in order to achieve its vision. Strategic Themes knit together independent activities, and focuses on effort and resources of functional groups that are significantly important. For instance, a budget airline's strategic theme can be "Now Everyone Can Fly", while a national airline's strategic themes can be "Your Best Travelling Experience".

Strategic Themes can be regarded as an organisation's pillar of excellence when they apply it to every part of the organisation, and define what major areas the organisation will focus to achieve its vision. This is illustrated in Figure 1.5.

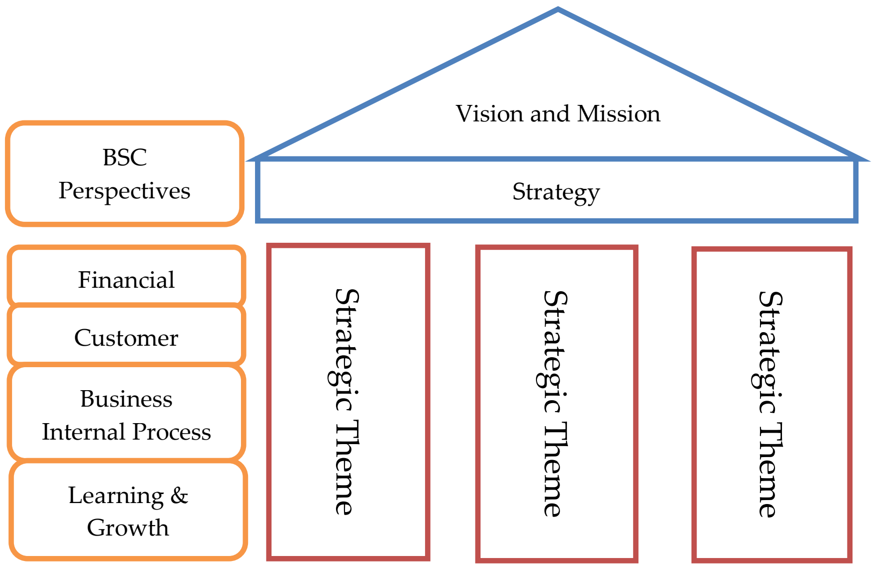


Figure 1.5: Strategic Themes as Pillars Supporting Organisation's Vision and Mission

Strategic Themes are intended to create shared understanding of what Strategic Objectives are, and actions to focus on (usually over a period of three to five years planning horizon). They should be designed to be measurable since they are considered to be something similar to major or top objectives. Furthermore, Strategic Themes can be used to separate an organisational strategy into several focused themes, and these themes can affect the four BSC perspectives.

Examples of a set of strategic themes for a leading fast-food-chain company include improving operating margin by re-engineering supply chain processes, and reducing cost of capital employed.

3.3. Strategic Objectives

Strategic Objectives are briefly defined objectives on the strategy map, which require further supporting information, initiatives and plans on how they can be achieved. Ideally, initial short-form wording of objectives should be meaningful, tangible, and written in everyday language that allows middle and lower level management personnel to understand the purpose of performing certain activities.

Performance management begins with Strategic Objectives. These objectives define the results that people aim to achieve. In the absence of Strategic Objectives, financial and human resources would be wasted on activities that contribute little towards organisational success. Strategic Objectives focus limited resources on things that matter most. Some objectives can be shared for better collaboration instead of individuals striking out in different directions.

Setting of Strategic Objectives is one of the most important management functions. A manager is responsible for setting objectives for a business unit and its members. The manager should also work with others to create plans that staff can follow to attain their goals.

Using the BSC as a platform, Strategic Objectives can be established for any broad strategy devised by an organisation. In the BSC model, each perspective has a defining question. For example, from the Financial Perspective, the defining question is: How should we appear to stakeholders?

In Figure 1.6, under the Financial Perspective, there are two Strategic Themes: Financial Growth and Productive Use of Financial Resources. These themes are translated into two strategic objectives, such as to increase revenue or to lower operations cost.

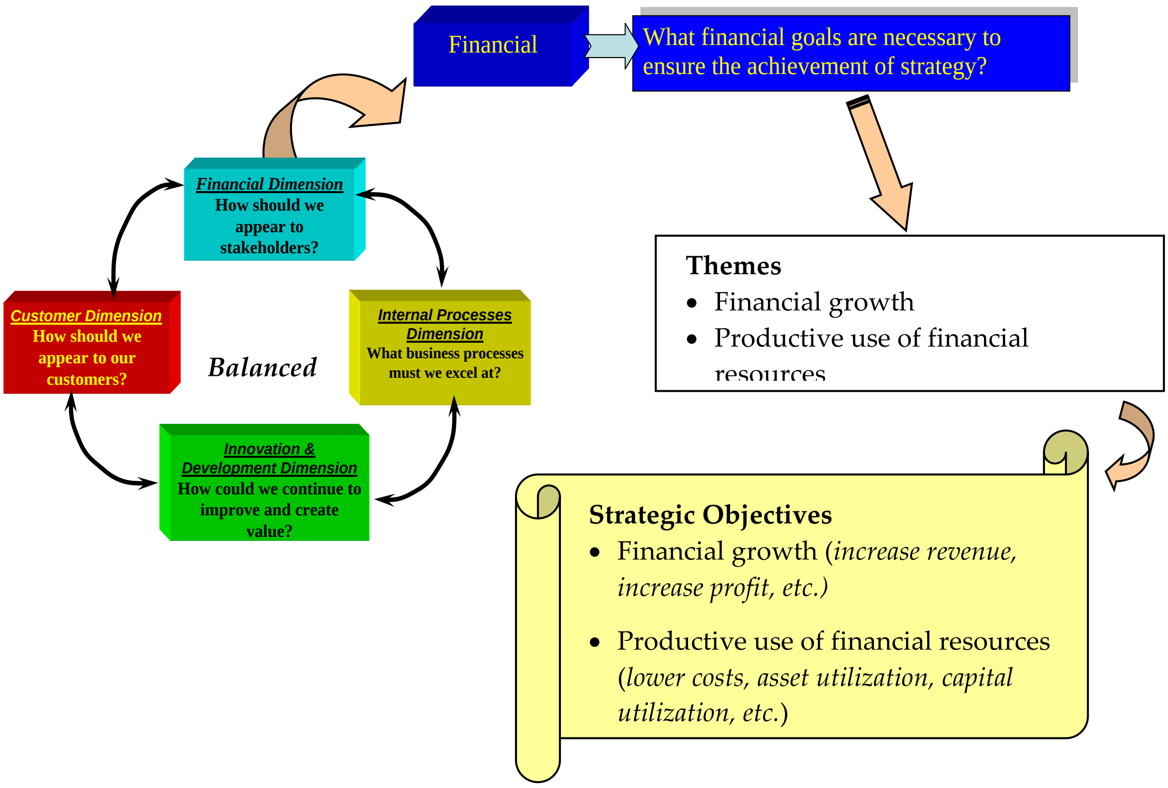


Figure 1.6: Strategy Themes to Strategic Objectives under BSC Financial Perspective

In order to monitor if the strategic objectives can be accomplished, the right Business Performance Measures must first be defined. Business Performance Measures are intended to provide objective information on the effectiveness and efficiency of a particular process or activity. This will lead to the accomplishment of goals. Linking a performance measure to a strategic objective provides relevance for the measurement. It also helps drive the organisation's activities towards achieving the organisation's strategy.

To select the appropriate measure, the following are considerations that need to be made are:

- whether the measure has an impact on the strategic objective,
- whether the measure is measurable,
- whether the data to derive the measure are easily available and accessible, and
- the measurement frequency.

3.4. Business Performance Measures

Business Performance Measures are the standards used to evaluate and communicate business performance against the expected Strategic Objectives. Well-defined Business Performance Measures not only provide managers and the top management with a tool to gauge organisational progress, they also inspire and motivate employees, set direction for the organisation, and encourage top-bottom alignment.

Although organisations may have long been devoted to the art and science of tracking business performance, it is common to see many organisations simply measure their business performance just to meet their internal audit requirements, instead of measuring the right aspects of their business performance.

Measures must be developed based on priorities of the strategic plan, which provide the key business drivers and criteria for measures that managers would like to see. Processes are then designed to collect data for these relevant measures with the objective of improving collection efficiency, storage and presentation.

Decision makers and managers rely on Business Performance Measures to examine their Strategy Objectives of various processes. The value of measures lies in their ability to provide a factual basis for defining:

- strategic feedback in order to show the present status of the organisation in achieving organisational Vision, Mission, and Strategy Objectives,
- diagnostic feedback to various processes to guide improvements on a continual basis,
- trends in business performance over time as the measures are tracked, and
- quantitative inputs to forecasting methods and models for decision support systems.

With reference to the theoretical foundations of Business Performance Measurement covered in Section 2.1.2, we can classify performance measures under the following categories:

1. **Input Measures**

   Input Measures track inputs such as production input, staff time, budgetary resources, etc. These are the simplest elements to measure, but offer limited information for decision making and analysis of actual results.

2. Output Measures

Outputs are results generated from the use of inputs. Some common Output Measures include tracking the number of people served, services provided, or units produced. Depending on the nature of the business activity and circumstances, Output Measures may provide information on whether the results are achieved.

3. Outcome Measures

Outcomes Measures track the benefits the stakeholders receive as a result of the organisation's activities or operations. Measuring an outcome shifts the focus from activities to results. Outcome Measures provide information on how an activity operates to the benefit of the stakeholders.

In designing the right measure, the following characteristics have to be carefully considered.

1. Is it measurable?

This relates to the objectivity of the measure and the ability of current technology to perform the measurement. Not all situations are directly measurable (e.g. customer or employee level of satisfaction). If no direct measure can be used, proxy measures are used in place. It is important that sufficiently accurate proxy measures are used to represent the direct measurement sought.

2. Are the data currently available or easily available?

Although a good measure can be selected after careful consideration, it may turn out that trying to collect data relating to the measure may be beyond the capability of the organisation at the current state of operations. Thus, data availability has to be factored into the selection of measures. Most of the time, business performance data are available, thus it is a matter of standardising the procedures to formally collect and collate data. However, if business performance data are not readily available, the question is then whether they can be made easily available. In such situations, new instruments may have to be used to generate the data. The instruments can be in the form of new sensors, new IT systems, or just a new survey questionnaires. It should be noted that when data are not easily available, considerable resources will have to be used to generate the data.

3. **What should the frequency be?**

   In determining the measurement frequency, it is important to select an appropriate time period for every measure. Measures associated with strategic performance tend to be measured on a quarterly or yearly basis, whereas measures associated with operational performance are more likely to be monthly, weekly, or even daily.

Once measures have been identified, proper documentation of selected measures must follow. The measure attributes to be defined and documented are

- the name of measure,
- relation of measure to a specific strategic goal,
- a specified target,
- formulae (if any) to compute values for the measure,
- frequency of data collection, and
- responsibility and ownership.

At times, it may be difficult to create a Business Performance Measure for business performance that is not easy to quantify, such as customer relations improvement. Although such measure is common and popular, measuring it is not always straight forward or possible. But we can measure components that contribute to customer relations improvement, such as the number of customer complaints satisfactorily handled during the month, or the average time required to resolve a customer's problem.

Likewise, we may encounter a situation where there may not be a single measure that properly or accurately represents the intent of a Strategy Objective. In such a circumstance, a special type of measure called *Proxy Measure* (or *Surrogate*) can be used. A Proxy Measure is an alternative choice of measurement used when a direct measure is unavailable. For example, a hospital assesses the effectiveness of medical facilities with a Proxy Measure like patients' average recovery time. Another example is the number of female members of a chamber of commerce as a Proxy Measure for estimating the percentage of female business owners.

A Proxy Measure is useful in providing a close approximate indication of actual business performance. Although it may not be perfect or conceptually ideal, it, nonetheless, can intuitively represent business performance until a better measure is available. A Proxy Measure is usually adopted when the data of the ideal measure is unavailable, or too expensive to collect, or when technically better measures are not intuitive to stakeholders. Proxy Measures are not usually used as long-term substitutes for ideal measures. However, an organisation will be better off using Proxy Measures than not measuring at all, until a better measure comes along.

The measures under the Balanced Scorecard approach should contain a mix of Leading and Lagging Indicators (measures) of business performance. If there is little or no leading indicators (i.e. leading indicator is essentially historical representation of business performance), we will end up knowing little about the "how" when it comes to operations. Conversely, a preoccupation with leading indicators will not reveal whether improvements are leading to improved process and results.

The differences between Leading and Lagging Indicators are summarised in Figure 1.7:

	Lagging	Leading
Definition	Measures focus on results at the end of a time period Based on historical performance	Measures that drive or lead to the performance of lag measures Measures intermediate process and activities
Examples	Revenue Employee Satisfaction	Punctuality Absenteeism
Advantages	Often easily identified and captured	Predictive in nature, and allow an organisation to make adjustments based on results
Issues	Historical in nature and do not reflect current activities Lack predictive power	May prove to be difficult to identify and capture. Often, new measures have no history
The Balanced Scorecard should contain a mix of leading and lagging indicators.		

Figure 1.7: Differences between Lagging and Leading Indicators (Measures)

The BSC-approach strategy maps are designed to help an organisation navigate its journey and the changing tides. As such, they must be updated to ensure that they remain relevant and effective. In today's environment, minor updates or modifications can be expected within 1-2 year's timeframe, while major changes occur when there is a change in business focus or strategy, resulting from a change of CEO or leadership.

3.5. Business Performance Targets

Once Business Performance Measures are selected, the next stage is to set Business Performance Targets that have to be achieved by an organisation. *Business Performance Targets* are objective values that share the same unit of measurement as the measures per se. *Targets* represent a performance level that the organisation wants to achieve in order to fulfil its stated strategy either partially or fully. With the identified targets, the gap between current and intended performance level can be ascertained.

Business Performance Targets give a sense of direction and purpose to establish Strategic Initiatives that will drive improvement activities and breakthroughs in the processes, and those activities that are being measured. Targets drive change by influencing behaviours, so people try to do the right thing the right way.

Business Performance Targets can be short-term (achieve within one year), mid-term (achieve within three years), or long term (achieve within five years). They usually correspond to an organisation's Strategic Objectives in terms of being incremental objectives, stretch objectives and visionary objectives respectively.

A systematic approach towards setting up targets can start by specifying the target for each strategic objective, and subsequently to each measure. Such an approach enhances the cause-and-effect linkages of the entire Business Performance Measurement system.

Targets should still be set even if managers are unsure about what they should be. Having an indicative target would still be preferred over a blank target, as it would drive certain business performance. In such situations, it is important to note and communicate appropriately that targets for such measures are merely indicated and subject to future refinement. Good commitment can also be obtained by getting relevant employees to participate in setting the future target.

Targets set must be reviewed on a regular basis to ensure the ongoing validity of the targets, and to ensure that employees are motivated towards achieving them. Targets that are too easily attained or too difficult to achieve will affect motivation of employees. Over time, the use of targets can form a basis for establishing internal benchmarks or best practices for improving internal processes.

The characteristics of a well-chosen target are:

1. easily understood and communicated
2. clear in establishing expectations
3. encouragements given to stretch performance

It is common to see organisations become too involved and mechanical in setting targets. They fail to realise that targets will shape organisation behaviours. Targets should be established to motivate people to perform in a certain way.

Poorly chosen targets can cause confusion, undermine strategy, cause people to pursue the wrong goals, unnecessarily consume resources, and de-motivate employees. As targets have cause-and-effect links to Strategic Objectives, it would certainly affect the overall strategy of the organisation.

At times, interim targets encourage an organisation to perform at a preferred pace. Progress towards a Strategic Objective is easily measured against multiple steps of the interim targets. Interim targets also offer a sense of progress and motivation.

3.6. Strategic Initiatives

Strategic Initiatives are specific projects, programmes or planned activities directed at key processes for the purpose of enhancing their output performance, and to meet or exceed established targets. Such initiatives are directed at closing the gaps identified during routine business performance measurement.

A gap occurs when the measured business performance falls short of the target. Therefore, in order to bring the level of performance up to a desirable level, corrective actions or improvements must be made to the processes. These improvement activities are incorporated as strategic initiatives.

3.7. Strategy Map

A *strategy map* is a diagram use to communicate the strategic objectives organisations would like to achieve. When we use a strategy map for the Balanced Scorecard approach, individual Strategic Objectives are positioned systematically to demonstrate the relationships between them, and reveal the logic under each Strategic Theme. A strategy map describes the process in transforming intangible assets into tangible customer and financial objectives. It also provides the management with a framework for describing and managing strategies in a knowledge-worker and strategy focused organisation.

Each Strategic Objective is related and linked to other Strategic Objectives on the strategy map using lines that represent cause-and-effect relationships. Such linkages logically reflect the natural value chain and culture. The strategy map also illustrates how the achievement of one objective enables the achievement of another, and captures the underlying hypotheses for the strategy.

The cause-and-effect relationships are critical components of a strategy map. They are represented by lines with arrows, linking various items in the strategic map such as Strategies, Strategic Themes, Strategic Objectives and Business Performance Measures.

Figure 1.9 illustrates the architecture of a strategy map for a balanced scorecard approach for a retail company specialising in men's apparels. The cause-and-effect logic of this design constitutes the hypotheses of the strategy. The financial perspective consists of two themes: *growth* and *productivity*, for improving shareholder value (i.e. profit). The themes in the customer perspective emphasise the importance of customer shopping experience, brand image, and product fit.

There are four strategic themes in the internal business process perspective: Brand Dominance, Fashion Excellence, Sourcing & Distribution Networks, and Customer Shopping Experience; to build operational excellence, increase customer values and build franchise. Finally, the four strategic themes in the Learning and Growth perspective are Strategic Awareness, Goal Alignment, Staff Competence, and Technology Infrastructure. They support the four internal business processes.

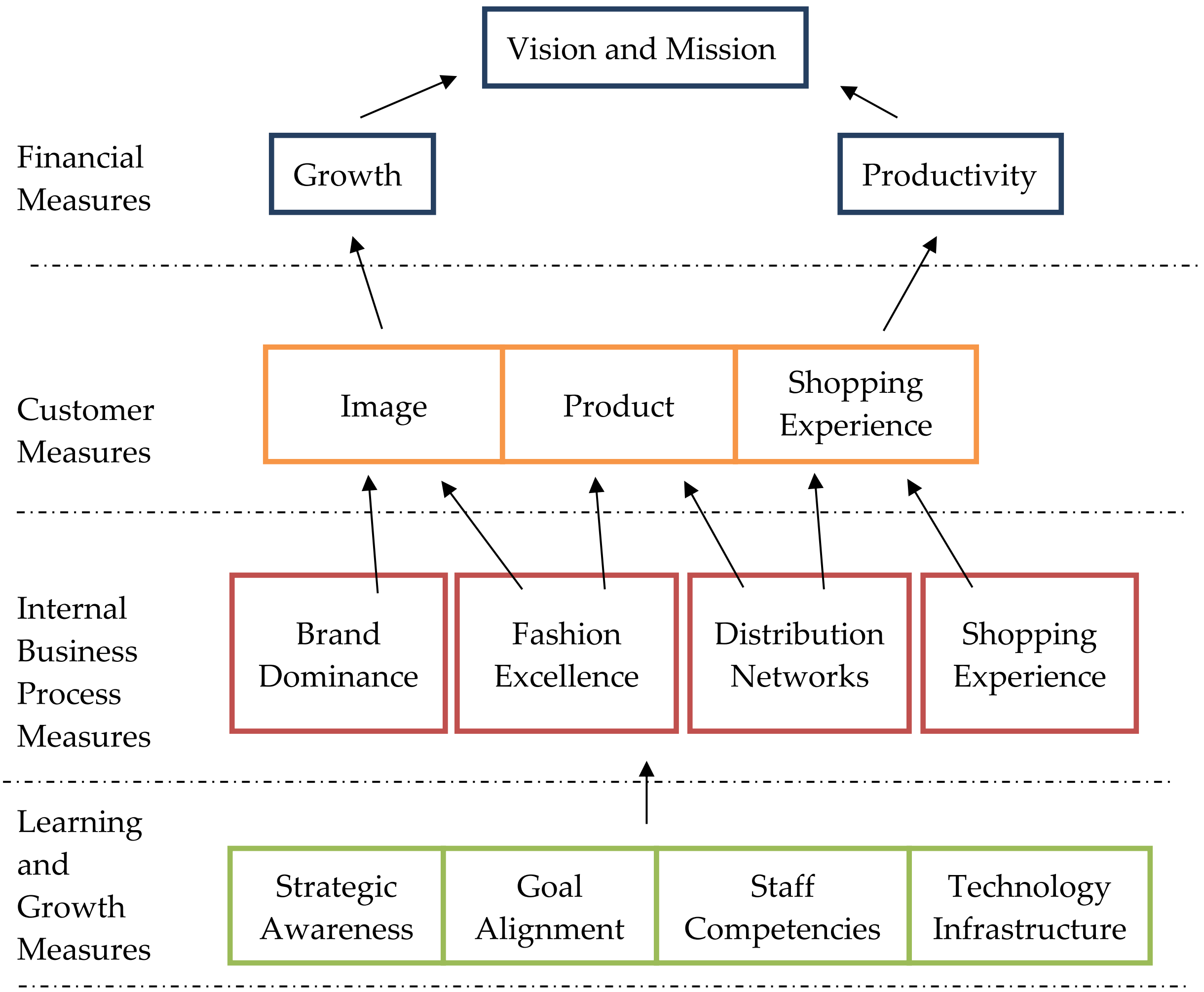


Figure 1.8: A Fashion Retailer's Balanced Scorecard Strategy Map

4. Internalisation of Business Performance Measurement

WATCH

Internalisation of Business Performance Measurement
(Access video via iStudyGuide)

The Business Performance Measurement holds tremendous promise when implemented with passion and commitment. Resistance to change should be replaced with pursuit for improvement and desire to advance towards the organisation's vision and mission.

An effective and sustainable Business Performance Measurement should be balanced by placing an appropriate balanced focus on the number of measures in all four BSC perspectives. There must be cause-and-effect relationships amongst Strategies, Strategic Objectives, Measures, and Targets.

To be truly aligned, an organisation must use the corporate strategy to determine how resources and processes can most effectively be used to achieve its goals. Accomplishing this requires the following steps:

1. Ensure all senior management staff reporting to the CEO or leadership has a clear understanding of the corporate strategy. It is common to see different Directors or Heads of Division having different interpretations of the corporate strategy. Such misinterpretations must be ironed out to ensure that all divisions' efforts are solely directed towards achieving a common objective.

2. Turn the corporate strategy into something actionable for staff. Action and action plan are critical, because without them many strategies and planning are just "paper exercise".

3. Implement procedures that would keep all business units aligned with the strategy. An action plan must be properly implemented and monitored. Over time, it is possible that certain business units become overly focused on their day-to-day activities that they begin to veer off track, and have their actions misaligned with achieving the corporate strategy.

4. A change in corporate strategy will affect many people in an organisation. The impact is most significant at the top level management, people on the ground

may not find it crucial. Therefore, it is important to create clarity around what the strategy means at the business unit, team and individual levels. To achieve this, the implementation team needs to involve key managers in defining how the strategy relates to their business units and what kind of alignment is required.

4.1. Obstacles in Implementing Business Performance Measurement Systems

Using the BSC approach to translate corporate strategies into actionable strategic goals that managers in the organisation can understand and act, is an efficient way to operationalise some of the principles of strategic Business Performance Management. There are obstacles to the successful implementation of Business Performance Measurement Systems as highlighted by Kaplan and Norton (2005):

- Lack of senior management support
- Too few individuals involved
- Too little consensus on measures
- Consensus sought only from the top management and not the others
- Keeping scorecard only at the top
- Prolonged development of the scorecard
- Scorecard that is alienated from the organisation strategy
- Treat the BSC as a system project
- Hire inexperienced Business Performance Management consultants

Implementation of a Business Performance Measurement System introduces something new that may not be immediately accepted by an organisation. It is critical to communicate (initial and on-going) with the employees regarding the new implementation. Time must be taken to demonstrate the new Business Performance System to employees, and explain to them how it can benefit them.

From the employees' point of view, it is certainly important for them to know sufficiently about the measure definitions, where the data come from, how the measure results are calculated, and how targets are developed. If these steps are not taken, employees may end up questioning results validity and simply remain uncooperative. Although there may be sceptics in the beginning, time and success stories can help overcome such hurdles.

4.2. Success Factors in Implementing and Sustaining Business Performance Measurement Systems

An often overlooked method that can enhance the implementation of a Business Performance Measurement System is advertising and celebrating success. The successful completion of a Business Performance Measurement System implementation stage or the attainment of an organisational objective is an ideal time to recognise it, and use such success as a "role model" for other business units, especially sceptics with negative views. Such effort can pique the interest of staff that has not yet been part of the process. Success often breeds success, and a success story can influence another business unit to follow the footsteps leading to actual success.

For example, an organisation can advertise a successful Business Performance Measurement System implementation of one department, division or business unit in an internal newsletter. Apart from advertising and promoting the initiative, such newsletter also educates the staff under business units that have not yet been involved, and encourages future involvement.

For organisations to implement and, more importantly, to sustain the operation of a Business Performance Measurement System, the following success factors are important considerations:

1. **Top Management Support**

   In order to provide a supportive and conducive environment for moving a strategy forward, the organisation must obtain buy-in from the management level top-down and bottom-up.

   Top management buy-in is important to the success of Business Performance Measurement System implementation. Such buy-in often drives acceptance of the system by employees at all levels of the organisation. Financial support and human resources will likely be diverted to a project like Business Performance Measurement System, only if the top management is convinced of its benefits. If the top management is not convinced, how will others in the organisation be convinced? Political battles can be minimised or avoided when the top management endorses and sponsors a Business Performance Management initiative, and publicises its benefits and importance to the organisation.

Bottom-up buy-in is also important because employees who are occupied with their daily work and responsibilities may not place much time and emphasis on a Business Performance Management project, which they regard as non-essential in fulfilling their job responsibilities according to their job descriptions.

However, having good buy-in does not mean that the Business Performance Management initiative will be spared from failure or termination. At times, a change in the top management can mean that the "pet projects" of the outgoing management team be eliminated.

To ensure the Business Performance Management initiative does not derail, it has to be integrated into the culture of the organisation. Using the results from Business Performance Management to set the agenda for routine management meetings or as key input to management decision process is essential to ensure survival and sustainability. Generally, management tools are not primary targets for termination when there is a change in management, but projects are, especially when there is a budget cut.

2. Project Champion

As with any other long running projects that involve the entire organisation, a project champion must be elected to drive the activities associated with Business Performance Measurement. The project champion should be someone from the senior management so as to be able to communicate effectively with all levels of the management.

3. Adequate Resources

Implementing a Business Performance Measurement System and then sustaining it over a long period of time requires a substantial amount of resources. Other than competent staff who track and collate data, the organisation has to deploy supportive IT systems and infrastructure to manage the collated data from Business Performance Measurements. These IT systems include automated data collection systems, data repositories, data warehouse, scalable server systems; and data analysis tools, such as business intelligence and analytics software. Compatibility issues must also be ironed out in order to facilitate the smooth transfer of data between systems.

4. **Employee Participation**

   To reduce resistance to change and obtain proactive participation, buy-in from the lower level management is required. In many situations, the measurement of business performance is undertaken at the source level, which usually involves a business process managed by employees within a particular department of an organisation. The responsibility of taking up Business Performance Measurements at regular periods will have to be left to the employees managing the process. Thus, it is imperative that employees accept this responsibility and undertake the Business Performance Measurement task conscientiously.

5. **Concise Reporting**

   Measurements collected have to be presented in a concise and efficient manner to managers, so that effective decision making can be done. Good reporting format and practices, like using performance dashboards, must be established.

6. **Linkage amongst Strategy, Strategic Themes, Strategic Objectives, Business Performance Measures, Business Performance Targets, and Strategic Initiatives**

   For a Business Performance Measurement System to be successfully implemented and useful, its Business Performance Measures must link to the organisation's strategies. Only then will the results from the measures be used to drive the organisation towards greater competitiveness, and profitability. Cause-and-effect linkage is one good method to link strategy to performance measures.

Developing a comprehensive Business Performance Measurement System is not cheap, simple or quick for the following reasons:

- Business performance is a multi-dimensional concept, which includes efficiency, effectiveness, quality, equity, fiscal stability, and conformance with government policy and standards.

- Data collection is not cheap and managers may not be able to capture all aspects of business performance.

- Those who will use business performance information may dictate which business performance dimensions to focus on.

- Although strategic objective plays a vital role in the designing of Business Performance Measurement System, the multiplicity of Strategic Objectives may exist and may be inconsistent at times.

- The lack of control of processes and time constraint may make measuring business performance difficult.

- Organisation barriers such as lack of accountability, insufficient analytical skills, lack of analytics staff, perverse reward system, inadequate performance evaluation and requirement of large investment upfront are possible hindrances to the implementation of Business Performance Management System.

- Personal barriers such as attitudes, traits, behaviours, risk avoidances, conceptual confusions, inadequate control, manager alibis, and perceived threats to job security are common barriers in implementing Business Performance Measurement System.

Summary

Most organisations today operate through decentralised business units and teams that are much closer to the customer than large corporate management staff. These organisations recognise that competitive advantage comes more from intangible knowledge, capabilities and relationships created by employees than from investments in physical assets and access to capital.

Strategy implementation, therefore, requires all business units, support units and employees to be aligned and linked to the strategy. With rapid changes in technologies, competition, and regulations, the formulation and implementation of strategies must become a continual and participative process. Organisations today need a language to communicate strategies, as well as processes and systems that help organisations to implement strategies and gain feedback about their strategies. Thus, success that comes from having a strategy becomes everyone's everyday job.

The Balanced Scorecard approach allows organisations to execute their strategies and offers many advantages to an organisation. First, it allows organisations to have a balanced view of its performance in all critical areas of the business. Second, it requires managers to first establish strategic objectives before the Business Performance Measures are selected. This ensures relevancy of the measures. Third, the Business Performance Measures channel improvement actions by the employees and the managers towards achieving the Strategic Objectives. Fourth, with balanced perspectives, selected measures tend to have a mixture of leading and lagging indicators. This ensures organisations are planning for the future, and at the same time understand the past and the current business performance.

Below is the summary on the process of linking strategies to Business Performance Measure.

1. Define Strategy: For example, to be the market leader by offering the best possible product to our customers.

2. Define Strategic Themes: For example, financial growth and productive use of financial resources.

3. Define Strategic Objectives: For example, under financial growth Strategic Theme, one of the Strategic Objectives can be identified as 50 percent increase in net profit in the next three years.

4. Define Business Performance Measures: For example, the percentage increment in net profit per year.

5. Conduct qualifying checks on the defined Business Performance Measures using the questions below:

   a. Does each measure have impact on the Strategic Objectives, and is it measurable?

   b. Are the data to derive the measures easily available to and accessible by the organisation?

   c. What should the measurement frequency be, and can the organisation update the data based on the preferred measurement frequency?

6. Establish a target for the measures: For example, the target percentage increment in net profit is equal to 20 percent per year.

A gap occurs when the measured business performance falls short of the target. In order to bring the level of performance up to a desirable level, corrective actions or improvements must be made to the processes. Such improvement activities are incorporated as strategic initiatives.

The Strategy Map offers a big picture in which Strategic Objectives are represented visually, and linked to one another through cause-and-effect relationships.

Identify the right Business Performance Measure is an important part in setting up a Business Performance Measurement System. If a performance cannot be measured, it is difficult to control and improve. Designing the right measure involves identifying the right type and attributes for the measure. Next, the right target must be set so that actual results can be compared against an appropriate target. Ownership must be assigned to every measure to ensure accountability.

The power and importance of alignment are not always evident in many organisations. Organisations need to align the cascaded scorecards with the higher-level scorecards to achieve the Strategic Objectives. While some Strategic Objectives and Business Performance Measures may be the same throughout an organisation, many will differ in each business unit in the organisation. Successful cascading is based on the clear understanding of the higher-level strategy map and positive influence by the lower-level contributors. All strategy maps must be verified to ensure true alignment.

Major obstacles that may affect the effective implementation of Business Performance Measurement System are discussed in the last section of this unit. The significance of this unit is to emphasise that Corporate Strategy, Business Performance Measurement, and continuous improvements are intertwined; all three are essential for an organisation to stay competitive. The top management buy-in is a critical step in attempting to obtain support and acceptance of the Business Performance Management by employees.

Exercises

ACTIVITY 1

Prepare your answers for the following questions as your instructor will ask you to present your answers in front of the class:

1. Why is Business Performance Management an important activity in an organisation?

2. List down five benefits of having Business Performance Management.

3. Is it sufficient just to focus on measuring the outcomes from an organisation? Why or why not?

4. Explain the differences between the past and the present Business Performance Measurement Systems.

5. Describe the four essential perspectives in the Balanced Score Card (BSC) Business Performance Management approach.

6. Describe the key components in a strategic Business Performance Management model.

7. Explain why it is important that an organisation's broad strategy has to be translated into Strategic Objectives.

8. You are a senior manager in a major company producing athletic shoes. Your organisation has just established the following two strategies:

- Build a strong momentum to grow the female consumer market.

- Drive increased margins through efficient inventory management.

Using the BSC approach, suggest appropriate Strategic Objectives and Business Performance Measures for your company.

9. Describe three obstacles that can impede the building of a Business Performance Measurement System.

10. Explain and elaborate what you understand by the statement: "If you can't measure it, you can't improve it."

11. Explain why choosing the right Business Performance Measure is important.

12. Are intangible Business Performance Measures identifiable? What should an implementer do if it is not identifiable?

13. Explain what the Leading and Lagging Indicators (measures) are.

14. Why is Business Performance Target important?

15. What are the consequences of having a Business Performance Measure without an owner?

16. What do you think are the most important success factors in implementing a sustainable Business Performance Measurement System?

ACTIVITY 2

1. Please read the Mobil Corporation's U.S. Marketing and Refining (USM&R) case study on http://www.jhtm.nl/ra/7114/artikel_mobil.pdf, and draw up Mobil USM&R's Strategy Map.

2. Form a group of four to five students. Discuss the strategy map your group is going to submit for course assessment.

References

Ariyachandra, T. R., & Frolick, M. N. (2008). *Critical Success Factors in Business Performance Management – Striving for Success.* Information Systems Management, Vol.25, 113-120.

Bacal, R. (1999). *Performance Management*. New York: McGraw Hill.

Harvard Business Review (1998). *Measuring Corporate Performance.* Boston: Harvard Business School Press.

Harvard Business School Press (2006). *Performance Management.* Boston: Harvard Business School Press.

Hobbs, L. (2008). *Strategic DNA.* Chicago: Agate Publishing.

Kaplan, R.S., & Norton, D.P. (2006). *Avoiding the Pitfalls.* Harvard Business Review, 2006, 1-14.

Kaplan, R.S., & Norton, D.P. (2008). *Mastering the Management System.* Harvard Business Review, January 2008, 1-7.

Kaplan, R.S. & Norton, D.P. (2005). *The Balanced Scorecard – Measures That Drive Performance.* Harvard Business Review, July-August 2005, 1-8.

Kaplan, R.S. (1996). *Mobil USM&R: Linking the Balanced Scorecard.* Boston: Harvard Business School Press.

Lawson, R., Desroches, D., & Hatch, T. (2008). *Scorecard Best Practices.* West Sussex: John Wiley & Sons.

Neely, A. (Ed) (2002). *Business Performance Measurement – Theory and Practice.* Cambridge: Cambridge University Press.

Neely, A., Bourne, M., Mills, J., Platts, K., & Richards, H. (2002). *Strategy and Performance: Getting the Measure of Your Business.* Cambridge: University Press.

Olve, Roy, & Wetter (1997). *Performance Drivers*: *A Practical Guide to using The Balanced Scorecard.* West Sussex: John Wiley & Sons.

Spitzer, D.R (2007). *Transforming Performance Measurement.* New York: AMACOM (American Management Association).

Tyagi, R. K., & Gupta, P. (2008). *A Complete and Balanced Service Scorecard.* New Jersey: FT Press.

STUDY UNIT 2
FOUNDATION OF THE SCIENCE OF DATA VISUALISATION

Learning Outcomes

At the end of this unit, you are expected to:

- describe what Data Visualisation is
- describe the benefits of Data Visualisation
- explain the four basic stages of Data Visualisation process
- illustrate how Data Visualisation can be used in everyday life
- describe what Semiotics of Data Visualisation is
- compare the properties of Sensory and Arbitrary Representations
- illustrate the concepts of Perceptual Processing Model
- describe what data can represent
- explain the two fundamental forms of data
- explain data attributes
- compare the four measurement levels of data quality attribute
- explain what metadata is
- prepare data using Data Visualisation software

Overview

This unit introduces the concepts, benefits and applications of Data Visualisation, the four basic stages of the Data Visualisation process, the Semiotics of Data Visualisation, and the Perceptual Processing Model used on Data Visualisation. In addition, this unit aims to discuss about what data can represent, two fundamental forms of data, data attributes, data quality measurement levels, and metadata.

1. Data Visualisation

WATCH

Data Visualisation
(Access video via iStudyGuide)

Data by definition means facts or information that are used to find out things or to make decisions (Oxford English Dictionary). Until recently, the term *Visualisation* meant the act of forming a picture of something or somebody in our mind (Oxford English Dictionary). However, *Visualisation* now means something more like a graphical representation of data or concepts. There is a shift on Visualisation's meaning from being an internal object in our mind to becoming an external object to support the decision-making process. Therefore, *Data Visualisation* means a process to create a visual representation of data.

1.1. Benefits of Data Visualisation

In this section, we will use an example to illustrate the benefits of Data Visualisation. Figure 2.1 shows a Data Visualisation of the Passamaquoddy Bay between Maine, in the United States; and New Brunswick, in Canada; where the tides are the highest in the world. Data come from approximately one million measurements and are collected by the University of New Brunswick's Ocean Mapping Group, using EM-1,000 multi-beam sonar system onboard the CHS Vessel F.G. Creed.

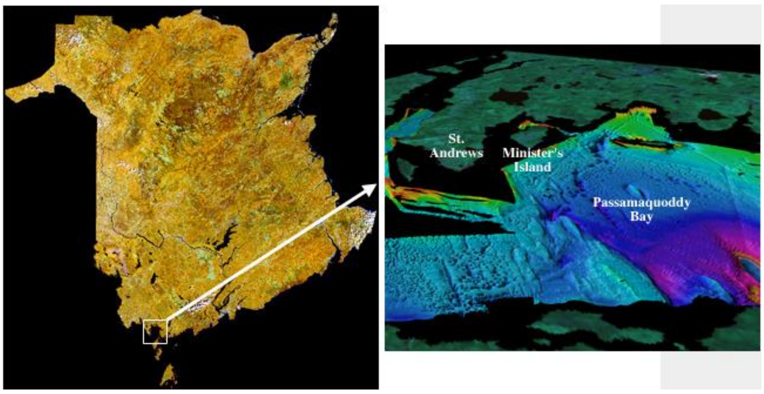


Figure 2.1: 3-D Data Visualisation of Colour-coded Bathymetry Data from Passamaquoddy Bay (University of New Brunswick, 2014)

If the above one million measurements data were presented using numbers rather than a 3D colour-coded map, it will be challenging for the research team from the University of New Brunswick to gain insight into the data.
The example on Passamaquoddy Bay highlights a number of benefits from Data Visualisation:

- Data Visualisation provides us the “ability” to comprehend huge amount of data. Thus, using Data Visualisation readers are able to interpret important information from more than a million measurements.

- Data Visualisation allows the perception of emergent properties that are not anticipated. The areas that have high tides (i.e. colour-coded in brighter colours) appear within the same cluster is immediately evident.

- Data Visualisation often enable problems with data to become immediately apparent. It is common for Data Visualisation to reveal things not only about the data, but also about the way the data are collected. With an appropriate Data Visualisation, missing data or errors in the data will surface. Thus, Data Visualisation can be invaluable in quality control.

- Data Visualisation facilitates the understanding of both large- and small-scale features of the data.

- Data Visualisation facilitates hypothesis formation. For example, the Data Visualisation in Figure 2.1 will help the research team from the University of New Brunswick in writing their research papers.

1.2. Data Visualisation Stages

According to Ware (2013), there are four basic stages in the Data Visualisation Process, together with three feedback loops as illustrated in Figure 2.2. The four basic stages are:

- Data Collection and Storage: the collection and storage of data

- Data Pre-processing: the pre-processing of data to transform them into something one can understand

- Graphics Engine: the display hardware and the graphics algorithms to produce Data Visualisation on screen

- Human Visual and Cognitive Processing: human perceptual and cognitive systems that are involved in interpreting and visualising data

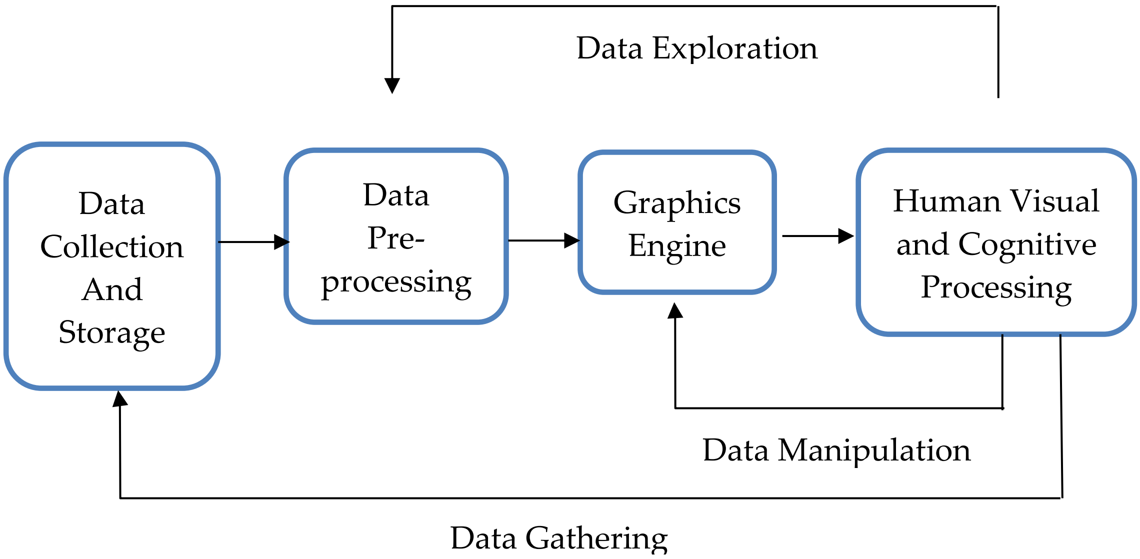


Figure 2.2: The Four Stages of Data Visualisation Process (Ware, 2013)

The longest feedback loop is the Data Gathering Loop where the data seeker, such as a business analyst or a market analyst, can choose to gather more data to follow up on an interesting insight provided by Data Visualisation. Both the physical and the social environments are involved in the Data Gathering Loop. The physical environment is the physical data sources, while the social environment determines in subtle and complex ways what is collected and how the data and Data Visualisation are interpreted.

The Data Exploration Loop controls the computational preprocess that takes place prior to Data Visualisation to transform the data before passing the data to the graphic engine.

Finally, in the Data Manipulation Loop, the Data Visualisation process may be highly interactive. For example, in a Data Visualisation, a market analyst may choose a different vantage point to better understand the market, where the analyst can use a laptop mouse to select parameter ranges that are most interesting.

1.3. Data Visualisation in Everyday Life

Data Visualisation has found its way into our everyday life. Almost all online contents are stored in databases and file systems. As people become more comfortable with computers and mobile devices, analysts and Data Visualisation software developers can create applications with interfaces that display the sheer quantity of data all at once.

1.3.1. Google News and Newsmap

Some of us perhaps like to use Google News to search for news. As shown in Figure 2.3, the Google News application shows a list of headlines, complemented with a thumbnail. Some of the top stories are listed at the top, while the most popular stories (i.e. based on a number of readers) are listed at the right sidebar. Other stories are categorised and shown below the top stories.

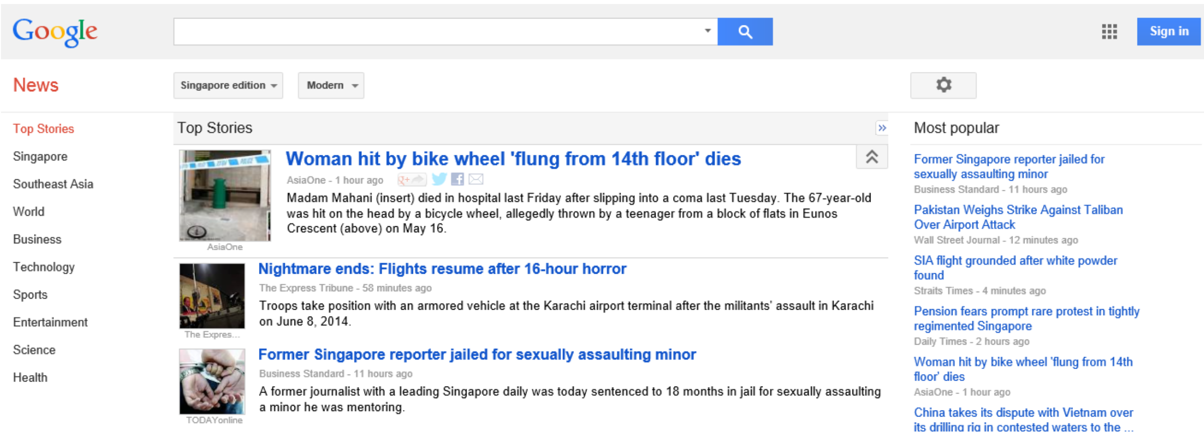


Figure 2.3: Google News (Google News, 2014)

If we were ever overwhelmed by the amount of news on Google News, the Newsmap Data Visualisation application that was launched in 2004 by Design Engineer Marcos Weskamp, offers us an option to view the news. The Newsmap application collects data from Google News and display the headlines sized by popularity (i.e. based on a number of readers), as shown in Figure 2.4.

Figure 2.4: Newsmap Data Visualisation application (Weskamp2014)

Each rectangle represents a "clickable story" and is coloured according to topics, such as world, national or business, so that users can get a sense of what is going on in the world at a glance. In addition, there is a variety of options, such as country of interest and time frame, for topics that users want to include and exclude as their profile preference.

1.3.2. Google Map and QQStay

Geographical maps are heavily used as a navigation tool where people can look up directions from Point A to Point B. The Google Maps application, as shown in Figure 2.5, was created to offer us navigation features with additional layer of information and in the context of areas, such as the nearby restaurants and other businesses.

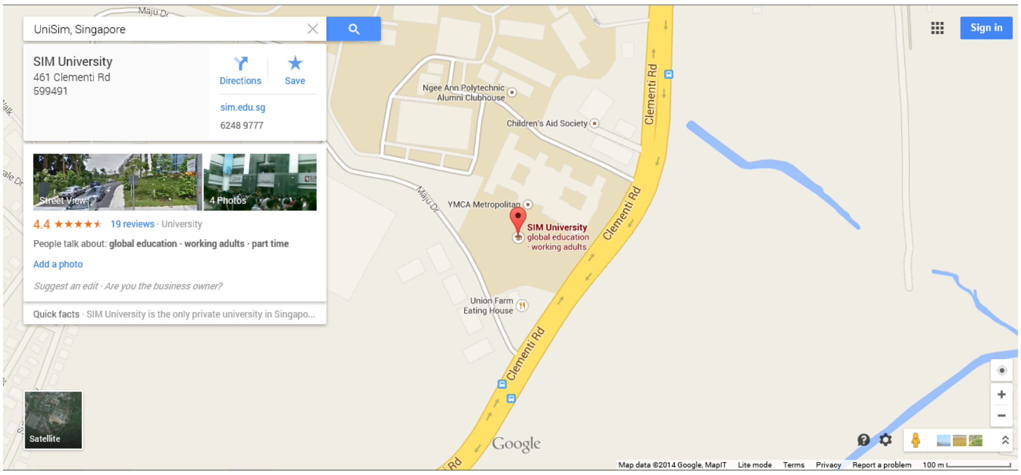


Figure 2.5: Google Map (Google Map on SIM University, 2014)

However, when we would like to know trends and patterns, or to get a general sense of an area, we can't have it on Google Maps because the application mostly shows only pointers and markers of specific locations.

QQStay, a web application that helps users who would like to search for real estate in Singapore, offers useful layers of information on property market trends. We can filter the property data by property type, prices, amenities, etc. Furthermore, the application is integrated with the social media platform, like Google+, Tweeter and Facebook, to send private messages to other users; and Google Calendar to save our home-viewing appointment, as shown in Figure 2.6.

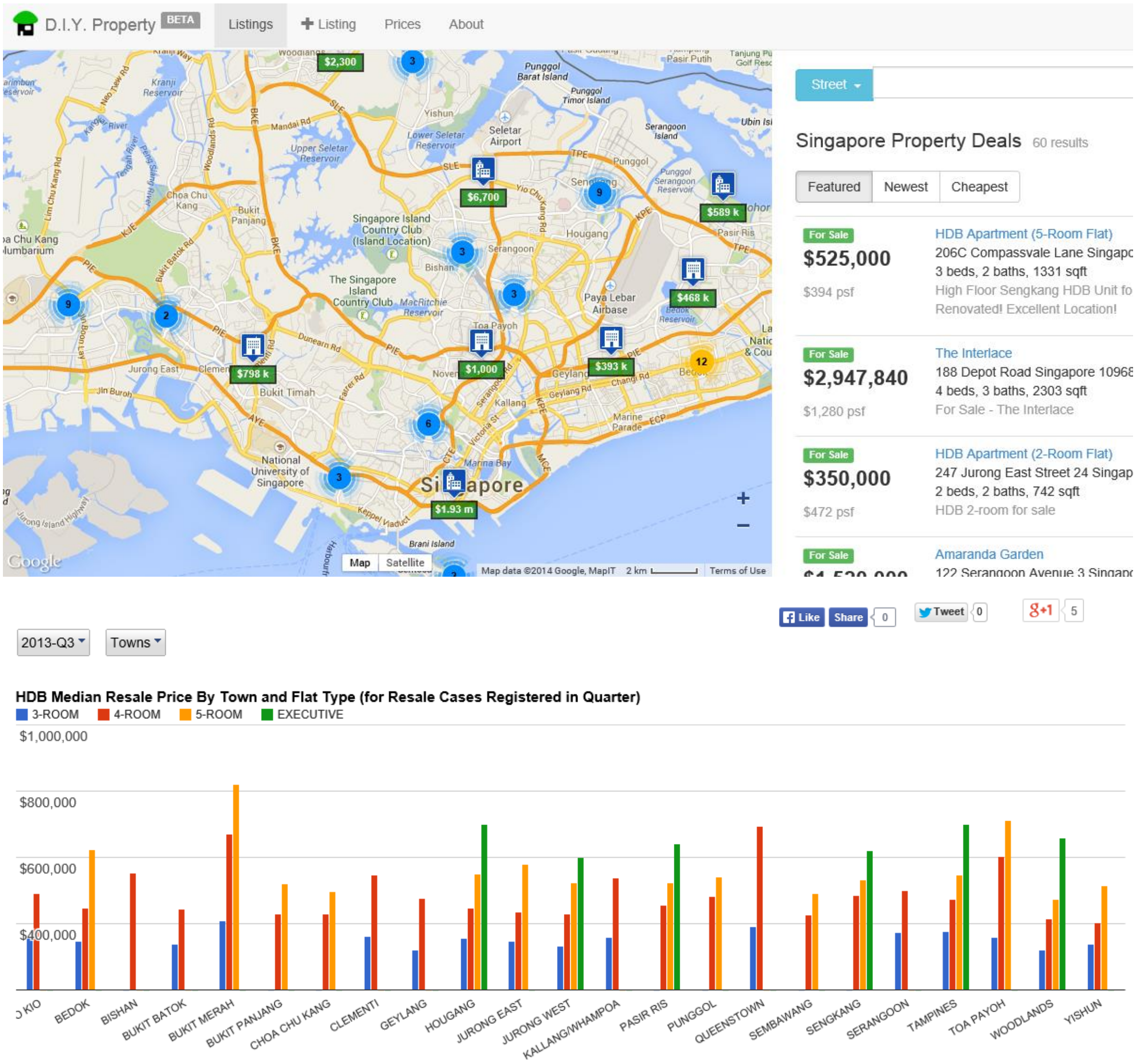


Figure 2.6: QQStay (QQStay, 2015)

1.3.3. Planetary

Some Data Visualisations may completely change the way we interact and relate to data. The Planetary (Bloom, 2014) is an iPad application that places iTunes music library into the context of a solar system, as shown in Figure 2.7.

Figure 2.7: The Planetary (Bloom, 2014)

This application uses stars to represent artists. Planets that orbits the stars represent albums, and moons that orbits the planets represent the music tracks. Instead of creating visualisation of opening a music library and choosing a specific song inside the music library, this application transforms visualisation into a landscape that we can explore and rediscover. Furthermore, as we use this application via the iPad touch interface, the data almost feels tangible to us.

2. Semiotics of Data Visualisation

Semiotics of Data Visualisation
(Access video via iStudyGuide)

Semiotics is the study of symbols and how they convey meaning. This discipline was originated in the United States by C. S. Peirce, and later developed in Europe by French philosopher and linguist Ferdinand de Saussure (see Saussure [1959]). Semiotics has mostly been dominated by philosophers, and by those who construct arguments based on examples rather than formal experiments. Bertin (1983) attempted to classify all graphical marks in terms of how they could express data. Most part of his work is based on his own judgement, though it is a highly trained and sensitive judgment. There are few, if any, references to theories of perception or scientific studies.

It is often claimed that visual languages are easy to learn and use, although in certain cases these languages are visual only to the extent that a written document is visual. It can be just as hard to learn to read some diagrams as it is to learn to read written language. Figure 2.8 shows three examples of languages that have some claims to be visual.

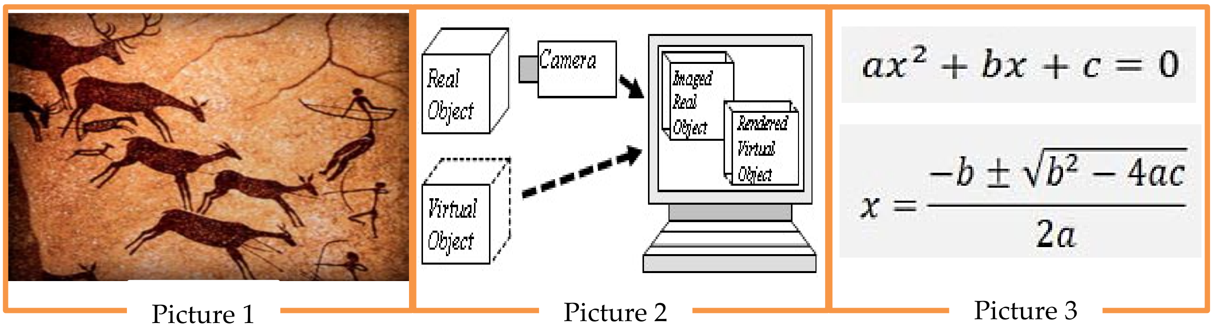


Figure 2.8: The three graphics and each of them can claim to be a visualisation. Picture 1 (Alvarado, 2013). Picture 2 (Freeman, 2005).

Picture 1 is a cave painting. We can interpret human figures and infer that in the figure, people are hunting deer with bows and arrows. Picture 2 is a schematic diagram of mixed realities where computer graphics (i.e. virtual object) are superimposed over the real world (i.e. real object). Picture 3 is a mathematical equation that may be totally unfamiliar to some people. The above pictures show that some visual languages are easier to read than others. We may think that it is because we have more experience with the kind of pictorial image shown in the cave painting, and less with schematic diagram of mixed realities, and mathematical equation.

Saussure originates the idea that there can be a science of visualisation. He defines a principle of arbitrariness, and applies it to the relationship between a symbol and the thing it signifies. Saussure was also a founding member of a group of structuralism philosophers and anthropologists who, although they disagreed on many fundamental issues, were unified in their general insistence that truth is relative to its social context.

Meanings to one culture may be nonsense to another as stated by thinkers such as Lévi—Strauss (1963a), Lacan (1964) and Barthes (1983). For example, a Tampines Expressway (TPE) or Central Expressway (CTE) road symbol in Singapore as a visual symbol for TPE or CTE can only be meaningful to people who drive or live in Singapore. Society creates meaning. We can interpret another culture only in the context of our own culture using our own language.

Languages are conventional means of communication in which the meanings of symbols are established through customs. Therefore, no one representation is better than another. All representations have value, and are meaningful to people who understand the representations and agree to the embedded meanings. It seems entirely reasonable to consider Data Visualisation as a form of communication, thus there can be a natural science of Data Visualisation.

2.1. Sensory versus Arbitrary Representation

"Realistic representation, in brief, depends not upon imitation or illusion or information but upon inculcation. Almost any picture may represent almost anything; that is, given picture and object there is usually a system of representation, a plan of correlation, under which the picture represents the object."

Nelson (1968)

The word *Sensory* refers to symbols and aspects of representation that derives its expressive power from its ability to use perceptual processing power of the brain without training. The word *Arbitrary* defines aspects of representation that must be trained, having no perceptual basis. For example, the written word "book" (i.e. arbitrary representation) has no perceptual relationship to any actual object.

Sensory representation can be understood without training. It can be processed rapidly and in parallel. It tends to be stable across individuals, cultures and time. Sensory representation is also resistant to instructional bias. Conversely, arbitrary representation is capable of rapid change and derives its power from culture. It can vary with culture and application.

As we understand the distinction between sensory and arbitrary representations, we nevertheless must recognise that most Data Visualisation are hybrids. In an obvious case, Data Visualisation may contain both pictures and words. But in many cases, sensory and arbitrary aspects of a representation are much more difficult to be separated. There is an intricate interweaving of learnt conventions and hard-wired processing in the human brain. Thus, the distinction between sensory and arbitrary representations is not as clear-cut as we would like them to be. We need to understand the properties of sensory and arbitrary representations to distinguish the distinction between sensory and arbitrary representations.

2.2. Properties of Sensory Representation

According to Ware (2013), below is the summary of the important properties of sensory representations.

2.2.1. Understanding without Training

A sensory representation is a visualisation which meaning is perceived without additional training. For example, it is immediately clear that the image in Figure 2.9 has several circles. Even though these circles are alien to most of us, the shape of the structure (i.e. circles) can be readily perceived.

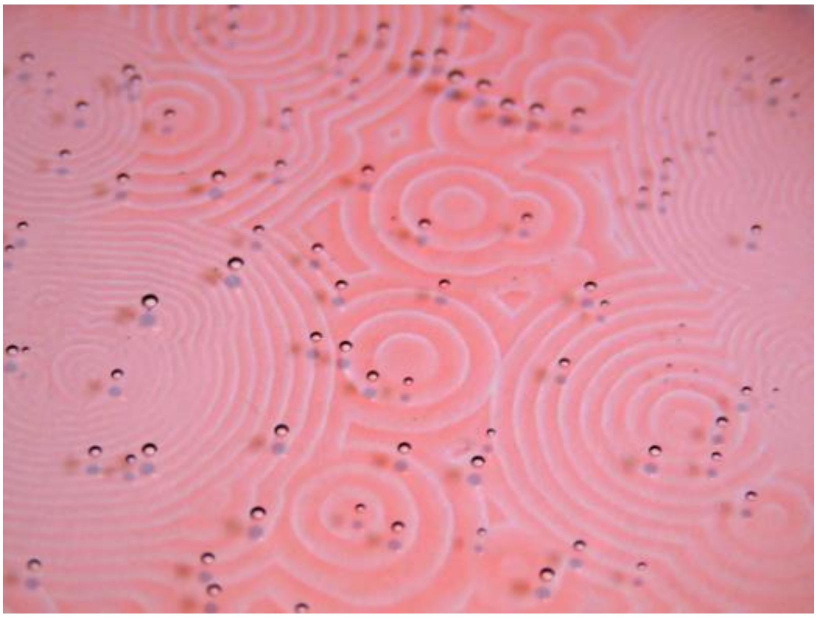


Figure 2.9: The concentric circles caused by a chemical reaction (Belousov-Zhabotinsky, 2014)

2.2.2. Resistance to Instructional Bias

Many sensory phenomena, such as the visualisations such as the one shown in Figure 2.10, persist despite the knowledge that those visualisations are illusory. When such illusions occur, they are likely to be misleading. But, what is important is that some aspects of the perception can be taken as the bottom-line facts that we ignore. In general, perceptual phenomena that persist and are highly resistant to change are likely to be hard-wired into the human brain.

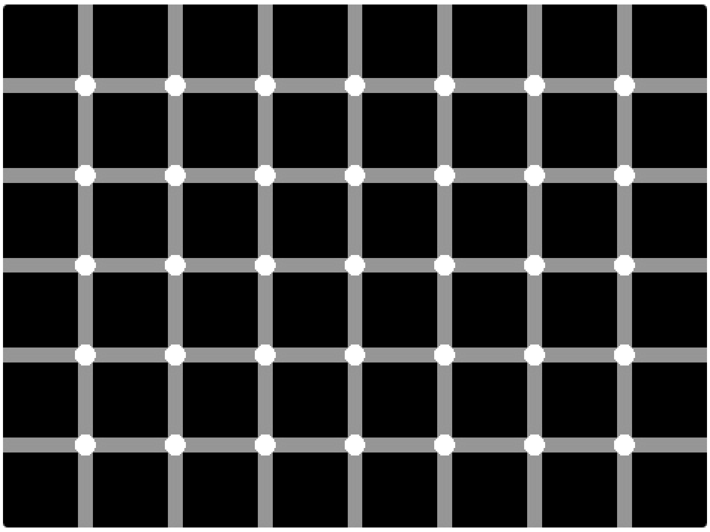


Figure 2.10: Black and White Dots Illusion (Pandey, 2013)

2.2.3. Sensory Immediacy

The processing of certain kinds of sensory information by the human brain is hard-wired and fast. Thus, we can code information in certain ways, so that the information will be processed by the human brain in parallel. This point is illustrated in Figure 2.11, which shows tilted Ts and upright Ts regions. The tilted Ts on the left appear to be a single patch. The region of tilted Ts is easy to differentiate from the neighbouring region of upright Ts. The way which the human visual system divides the visual world into regions is called *Segmentation*. This evidence suggests that this is a function of early rapid-processing systems in the human brain.

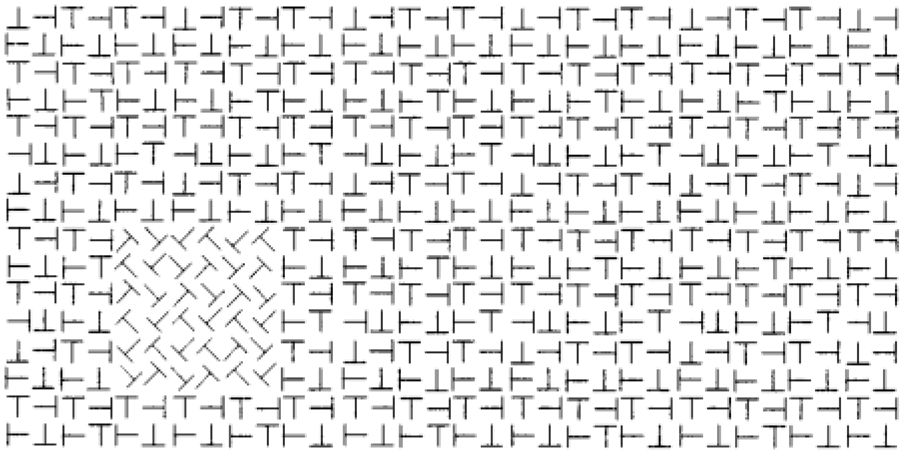


Figure 2.11: Tilted Ts and Upright Ts (Tomonaga, 1998)

2.2.4. Cross-cultural Validity

In general, a sensory representation can be understood across cultural boundaries. These may be national boundaries, or the boundaries between different user groups. Instances in which a sensory representation is misunderstood occur when some groups dictate that a sensory representation is to be used arbitrarily in contradiction to natural interpretation. In this case, the natural response to a particular patter will, in fact, be wrong.

2.3. Properties of Arbitrary Representation

According to Ware (2013), below is the summary of some basic properties of arbitrary representations.

2.3.1. Hard to Learn

It takes a child hundreds of hours to learn to read and write, even if the child has already acquired the spoken language. Graphical representations of alphabets and their rules of combination must be laboriously learnt. The Chinese character set is reputed to be even harder to work with than the Roman's.

2.3.2. Easy to Forget

Arbitrary representation that is not overlearned can easily be forgotten. It is also the case that arbitrary representation can interfere with each other. In contrast, the sensory representation cannot be easily forgotten because they are hard-wired, and forgetting them would be like learning not to see. Still, some arbitrary representations, such as written numbers, are overlearned to the extent that they will never be forgotten. Thus, we cannot always choose to use the most easily perceived display solution.

2.3.3. Embedded in Culture and Applications

In most of the Asian culture, red is normally used as a warning, while green symbolises renewal. However, in China, green symbolises death, while red symbolises good luck and fortune. The use of colour codes to indicate meanings is highly culture-specific.

Many graphical symbols are transient and tied to a local culture or application. Culturally embedded aspects of visualisations persist, because they have become embedded in ways in which we think about problems. For many geologists, the topographic contour map as shown in Figure 2.12 is the ideal way to understand relevant features of the earth's surface. They often resist shaded computer graphic representations, even though these appear to be much more intuitively understandable to most people. Contour maps are embedded in cartographic culture and training.

Figure 2.12: Contour map of Bryce Canyon National Park, Utah, the U.S.A (Creative Commons Attribution-ShareAlike, 2014)

2.3.4. Formally Powerful

Arbitrary representations can be constructed to embody formally defined languages. For example, Mathematicians have created hundreds of equations and graphical languages to express and communicate their concepts. Although the expressive power of mathematics in conveying abstract concepts in a way that is formal and rigorous is unequalled, the language of mathematics is extremely hard to grasp for most people. This is clear evidence that visual representation does not always mean that it will be easy to understand.

2.3.5. Capable of Rapid Change

One way of looking at the distinction between sensory and arbitrary representations is the development time of the representations. Sensory representations are the results of million-year evolution of our visual systems. Although arbitrary representation's evolution time frames are much shorter than sensory representation's time frames, they still lasted for thousands of years. With the help of high-performance interactive computer graphic technologies, we have capability to create new visualisations. We can now control motions and colours with great flexibility and precision. Because of the advancement in technologies, we are now witnessing an explosive growth in the invention of new visualisations.

2.4. Perceptual Processing Model

Figure 2.13 shows the information processing model of human visual perception according to Ware (2013).

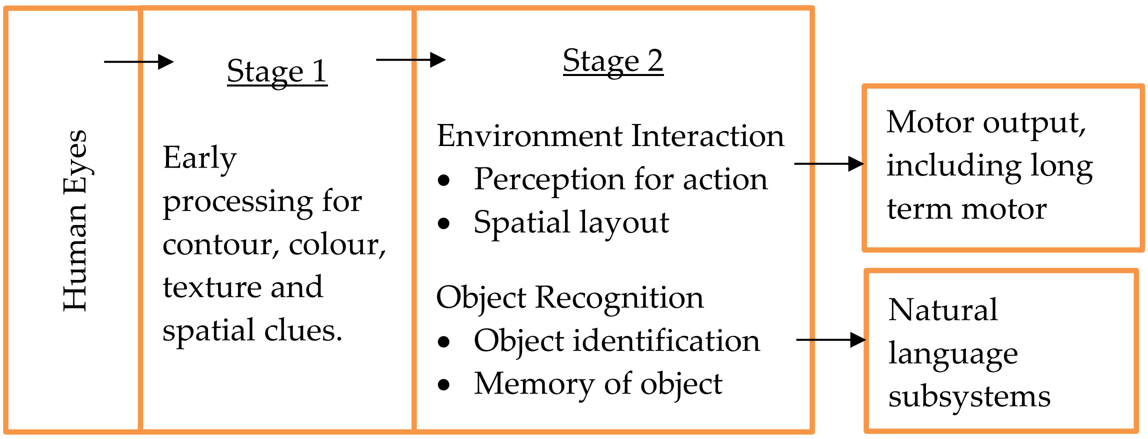


Figure 2.13: Overview of a two-stage model of human visual information processing (Ware, 2013)

In the first stage, information is processed in parallel to extract basic features of the environment. In the second stage, visual attention plays a much more active role, and items in the environment tend to be examined sequentially.

Stage 1: Parallel Processing to Extract Low-level Properties of the Visual Scene

Visual information is first processed by large arrays of neurons in the eye, and in the primary visual cortex at the back of the brain. Individual neuron is selectively tuned to certain kinds of information, such as the orientation of edges or the colour of a patch of light. In each subarea, large arrays of neurons work in parallel, extracting particular features of the environment. In the early stages, this rapid parallel processing proceeds whether we like it or not. It is largely independent of what we choose to attend to, not where we look at. If we want people to understand information quickly, we should present it in such a way that it can easily be detected by these large, fast computational systems of the brain.

Stage 2: Sequential Goal-directed Processing

The second stage is divided into two separate parts – a subsystem specialised in object recognition, and a subsystem specialised in interacting with the environment. To identify an object, people must match the object's visual characteristics with properties of the object stored in their memories. In addition, the task that the observer is performing will also influence what is perceived. One of the primary mechanisms relating what is perceived to the task is visual attention. Some aspects of the second stage processing occur sequentially where one visual object is processed at one time.

There is increasing evidence that tasks involving eye-hand coordination and locomotion may be processed in pathways distinct from those involved in object recognition. Milner and Goodale (1995) suggested two-visual-system hypotheses: one system for locomotion and action, another for symbolic object manipulation.

A number of theorists suggests that there is an intermediate representation of the world at the boundary between Stage 1 and Stage 2 processing. Marr (1982) called this the $2^1/_2$ sketch. Triesman (1980) called it a *feature map*. But much is still unknown about how visual objects are constructed from features, memories and the locus of visual attention.

Beyond the visual processing stages shown in Figure 2.13 are interfaces of other subsystems. A visual object recognition process interfaces with the verbal linguistic subsystems of the brain, so that words can be connected to images. The environment interaction subsystem interfaces with the motor systems that control muscle movements.

3. Understanding Data

WATCH

Understanding Data
(Access video via iStudyGuide)

Data are more than just numbers. To visualise data, we must know what the data represent. Data represent real life. A single data point can have who, what, when, where and why attached to it. Extracting information from a data point is challenging. We can make assumptions about data, such as how accurate it is or how it relates to its surroundings, and end up with a skewed view of what the data actually represents. Thus, we need to look at everything around us, identify the context and see what the data set looks like as a whole.

It is often impossible to count or record everything because of cost and/or lack of manpower. We take bits and pieces, and look for patterns and connections to make an educated guess about what the data represent. Data is a simplification or abstraction of the world. When we visualise data, we visualise an abstraction of the world. Visualisation is an abstraction of data, so in the end, we end up with an abstraction of an abstraction.

3.1. Fundamental Forms of Data

The goal of visualisation is to transform data into a perceptually efficient visual format. Bertin (1977) suggested that there are two fundamental forms of data, which are data values and data structures. A modern way of expressing this idea is to divide data into *Entities* and *Relationships* (often called Relations).

Entities are the objects we wish to visualise, while Relationships define the structures and patterns that relate Entities with one another. Sometimes, the relationships are provided explicitly, and discovering relationships is the very purpose of visualisation.

3.1.1. Entities

Entities are generally objects of interest, for example, products and sales persons. A group of objects can be considered as a single entity by data visualisation designers. For example a number of products can be considered as a single product category entity, and a number of sales persons stationed in Singapore can be considered as a single office entity.

3.1.2. Relationships

Relationships form the structures that relate to entities. There can be many kinds of relationships. For example, a laptop screen has a "part-of" relationship to a laptop product, and one sales person of a firm may have a "supervisory" relationship with one sales manager of a firm. Relationships can be structural and physical, as in defining the way a product is made of its many components or parts; or they can be conceptual, as in defining the relationship between a sales person and its sales manager. Relationships can also be causal, as when the Malaysian Ringgit depreciation event in 2015 causes the influx of tourists in Malaysia; or they can be purely temporal, defining an interval between two events.

3.2. Data Attributes

Both Entities and Relationships can have attributes. In general, something should be called an *Attribute* when it is a property of some entities and cannot be thought of independently. For example, the colour of a car is an attribute of the car, the taste of food is an attribute of the food, and the floor area is an attribute of a house.

Defining what should be an entity and what should be an attribute is not always straight forward. For example, the sales price of a laptop could be thought of as an attribute of the laptop, but we can also think of the sum-of-money as an entity in itself. In this case we have to define the relationship between the laptop entity and the sum-of-money entity.

3.2.1. Attribute Quality

It is often useful to describe Data Visualisation methods in light of the quality of attributes they are capable of conveying. A useful way of considering the quality of data is the taxonomy of number measurements defined by the statistician S. S. Stevens. According to Stevens (1946), there are four levels of measurement:

1. ***Nominal***

   Nominal measurement measures items based on their labels or categories or other qualitative classification the items belong to with no implied order. Two types of nominal assignments are Type A where the "numbering" of product represents the identification of the product, and Type B where the "numbering" for each member of a same product type or product category is assigned the same numeral. Since the purpose is served when any two designating numerals are interchanged, this measurement form remains invariant under the general substitution or permutation group (sometimes called the *symmetric group of transformations*).

The only statistic relevant to nominal measurements of Type A is the number of cases. For example, the count of product identification numbers in the company's sales order system. For Type B, we can determine the most common item (the mode), and under certain conditions we can test, through contingency methods, hypotheses regarding the distribution of cases among the classes, for example the most frequent product categories bought by customers or the count of products in each product categories.

2. *Ordinal*

The ordinal measurement arises from the operation of rank ordering. An example of an ordinal measurement is the measurement of customer satisfaction rating. Other instances are found among measurements of examination grade and quality of product, or measurements used by psychologists.

"In the strictest propriety, the ordinary statistics involving means and standard deviations ought not to be used as the ordinal measurements, because these statistics imply a knowledge of something more than the relative rank-order of data. Means and standard deviations computed on the relative rank-order data are in error to the extent that the successive intervals on the measurements are unequal in size. When only the rank-order of data is known, we should proceed cautiously with our statistics, especially with the conclusions we draw from them.

Even in applying those statistics that are normally appropriate for the rank-order data, we sometimes find rigor compromised. Thus, although it is indicated in Figure 2.14 that the percentile measures may be applied to the rank-ordered data, it should be pointed out that the customary procedure of assigning a value to a percentile by interpolating linearly within a class interval is, in all strictness, wholly out of bounds. Likewise, it is not strictly proper to determine the mid-point of a class interval by linear interpolation, because the linearity of an ordinal measurement is precisely the property which is open to questions. "

Stevens (1946)

3. *Interval*

The Interval measurement allows us to measure the degree of difference between items, but not the ratio between them. We can use almost all the usual statistical measurements here, but we cannot use the interval measurement to measure items that have a zero point.

An example of the interval measurement is the measurement of temperature with the Centigrade or Celsius scale. The Celsius scale defines the freezing and boiling points of water at specific conditions, then separate both points into 100 intervals. Ten degree Celsius cannot be said to be thrice as hot as 30 degree Celsius. However, one temperature difference can be said to be thrice of another.

Other example is the measurement of time. When we measure date from an arbitrary epoch like AD, the measurement can be transformed to those on another by using $y=ax+b$ equation, but it will be meaningless to say that one value in an arbitrary epoch is twice or some other proportion greater than value in another arbitrary epoch. Periods of time, however, can be measured using ratio measurements, because one period may be correctly defined as double of another.

4. *Ratio*

Ratio measurements estimate the ratio between a magnitude of a continuous quantity and a unit magnitude of the same kind (Michell, 1997, 1999). Under ratio measurement, we can use the operations for determining all four relations: equality, rank-order, equality of intervals, and equality of ratios. An absolute zero is always implied, even though the zero value on some scales (e.g. absolute temperature) may never be produced. All types of statistical measures are applicable to ratio measurements.

Level of Measurement	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics
Nominal	Determination of equality	Permutation group x′ = f(x) f(x) means any one-to-one substitution	• Number of cases • Mode • Contingency correlation
Ordinal	Determination of greater or less	Isotonic group x′ = f(x) f(x) means any monotonic increasing function	• Median • Percentiles
Interval	Determination of equality of intervals or differences	General linear group x′ = ax + b	• Mean • Standard deviation • Rank-order correlation • Product-moment correlation
Ratio	Determination of equality of ratios	Similarity group x′ = ax	• Coefficient of variation

Figure 2.14: Overview of the four levels of measurement (Stevens, 1946)

When we need to decide the kind of measurement to be used, we may want to find out in what ways we can transform the values of items that we want to measure, and still have the values serve all the functions previously fulfilled. For a ratio measurement, the values of the items that we want to measure can be multiplied by a constant, which changes the size of the unit. A constant can also be added to the values (or a new zero point chosen). If the purpose of the measurement is still served when the values of the items that we want to measure are squared or cubed, it is not an interval measurement. Finally, if any two values may be interchanged at will, the ordinal measurement is ruled out and the nominal measurement is the sole remaining possibility.

In practice, only three of Stevens' level of measurement are widely used, in somewhat different form. The typical basic data classes most often considered in Data Visualisation have been greatly influenced by computer programming. They are as follows:

- Category Data: this is like Stevens' nominal measurement.

- Integer Data: this is like Stevens' ordinal measurement. It is discrete and ordered.

- Real-number Data: this combines the properties of Stevens' interval and ratio measurement.

The three data classes above can be very useful in discussing the expressive power of Data Visualisation techniques. For example, here are two generalisations:

- Use a graphic size, as in a pie chart, to display category data is likely to be misleading, because we tend to interpret size as representing quantity.
- Perceive nominal and ordinal values using colours. However, it is not very effective to perceive interval measurement using colours.

3.2.2. Attribute Dimension

An attribute of an entity can be a single dimension or multiple dimensional attribute. We can also have a *scalar* attribute quantity, such as the height of a building; or a *vector* attribute quantity, such as the direction in which a bus is travelling. *Tensor* is higher-order attribute quantity that describes both direction and shear forces occur in materials that are being stressed.

Sometimes, we can also have a *field of scalar, vector or tensor*. For example, the gravitational field of a planet is a three-dimensional vector field attribute. If we are interested only in the strength of gravity at the planet's surface. It is a two-dimensional scalar attribute. Often, the term *map* is used to describe this kind of field, like gravity map or temperature map.

3.3. Metadata

"When we strive to understand data, we may discover the correlation among variables or clusters of data values. We may also postulate certain underlying mechanism that are not immediately visible. The result is that theoretical entities, like atoms, photons, black holes, and all other basic constructs of physics, come into being.

As more evidence accumulates, theoretical entities seem more and more real, but they are nonetheless only observable in the most indirect ways. These theoretical constructs that emerge from data analysis are called *Metadata*."

(Tweedie, 1997)

Metadata are also known as *Derived Data* in the database modelling community because they are data about data. Metadata are structured information that explain, describe or locate the original (i.e. also known as primary data), otherwise make the using of original data more efficient. The comparison between primary data and their metadata are illustrated in Figure 2.15.

Bloomberg's U.S. Indexes data

Index	Value	Change	% Change	1 Month	1 Year	Time
Dow Jones Industrial Ave...	16,947.08	+25.62	+0.15%	+2.05%	+14.51%	16:15:00
S&P 500 Index	1,962.87	+3.39	+0.17%	+3.28%	+23.26%	16:32:38
NASDAQ Composite Index	4,368.04	+8.71	+0.20%	+4.35%	+30.11%	17:16:00
New York Stock Exchange ...	11,018.11	+15.03	+0.14%	+3.15%	+22.17%	20:59:48
S&P/TSX Composite Index	15,108.97	-3.25	-0.02%	+2.73%	+25.95%	18:05:00
Mexican Stock Exchange M...	42,865.72	-91.64	-0.21%	+2.26%	+12.70%	16:06:27
Mexican Stock Exchange I...	2,553.38	-7.68	-0.30%	+2.10%	+16.80%	16:06:27
Ibovespa Brasil Sao Paul...	54,638.19	-564.35	-1.02%	+3.82%	+16.11%	16:16:32
Santiago Stock Exchange ...	3,870.20	-23.36	-0.60%	-2.27%	+1.27%	16:36:03
Mid Market Index	137.82	+0.46	+0.33%	+4.57%	+22.07%	17:00:00

Bloomberg's U.S. Indexes Metadata

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8"/>

    <meta http-equiv="X-UA-Compatible" content="IE=10; IE=9; IE=8"/><script type="text/javascript">var
NREUMQ=NREUMQ||[];NREUMQ.push(["mark","firstbyte",new Date().getTime()]);</script>
<title>Stock Market Overview & Stock Indexes - Bloomberg</title>
    <meta name="keywords" content="stocks, major stocks, stock market news, stock indexes, stock indices, stock markets, top gainers,
top losers, world indexes, world future indexes, movers by index, stock market, stock indexes, stock exchanges" />

<meta name="description" content="Updated stock market overview with current price and changes. Find top gainers & losers by index,
major world indexes, futures indexes and stock news." />
<meta property="fb:admins" content="1502271052,1617206,500566238,616620419,18776,1396618572,501161086,100001111898866,1223836"/>
<meta property="og:title" content="Stock Market Overview & Stock Indexes"/>
<meta property="og:site_name" content="Bloomberg"/>
<meta property="og:description" content="Updated stock market overview with current price and changes. Find top gainers & losers by
index, major world indexes, futures indexes and stock news." />

<meta name="twitter:card" content="summary"/>
<meta name="twitter:site" content="@BloombergNews"/>
<meta name="twitter:creator" content="@BloombergNews"/>
<meta property="og:url" content="http://www.bloomberg.com/markets/stocks/"/>
<link rel="shortcut icon" href="/favicon.ico" type="image/x-icon" />


            <link href="http://cdn.gotraffic.net/v/20140620_134838/stylesheets/main-min-v3_1.css" media="all" rel="stylesheet"
type="text/css" />        <link href="http://cdn.gotraffic.net/v/20140620_134838/stylesheets/main-min-v3_2.css" media="all"
rel="stylesheet" type="text/css" />        <link href="http://cdn.gotraffic.net/v/20140620_
134838/stylesheets/compiled/reskin/site.css" media="all" rel="stylesheet" type="text/css" />
```

Figure 2.15: Example of Data versus Metadata (Bloomberg, 2014)

Summary

Data Visualisation applies visual research to practical problems of data analysis in much the same way as engineering applies physics to practical problems of building manufacturing plants. Just as engineering has influenced physicists to become more concerned in areas such as semiconductor technology, we may hope that the development of an applied discipline of Data Visualisation can encourage visual researchers to intensify their efforts in addressing problems, such as task-oriented perception and user interface in 3D space.

In this unit, we have introduced the key distinction between the ideas of sensory and arbitrary representations. This is a difficult and sometimes artificial distinction. We can come up with counter examples and reasons why it is impossible to separate the two. Nonetheless, the distinction is essential. Without a basic model of visual processing on which we can support the idea of a good data representation, the problem of Data Visualisation ultimately boils down to establishing a consistent representation. If the best representation is simply the one we know best, because it is embedded in our culture, then standardisation is everything – there will be no good representation, only widely shared conventions.

In opposition to the view that everything is arbitrary, we view all humans have more or less the same visual system. This visual system has evolved over tens of millions of years to enable humans to perceive and act within the natural environment. Although very flexible, human visual system is tuned to receiving data presented in certain ways. If we can understand how the mechanism works, we will be able to produce better Data Visualisation.

Exercises

ACTIVITY 1

Prepare your answers for the following questions as your instructor will ask you to present your answers in front of the class:

1. Explain the definition of Data Visualisation.

2. Describe the benefits of Data Visualisation.

3. Explain the four basic stages of the Data Visualisation process.

4. Illustrate how Data Visualisation can be used in our everyday life.

5. Explain the definition of the semiotics of Data Visualisation.

6. Compare the properties of Sensory and Arbitrary representations by explaining the similarities and the differences between the properties of Sensory and Arbitrary representations.

7. Illustrate the concepts of Perceptual Processing Model.

8. Describe what data can represent our everyday life.

9. Explain the two fundamental forms of data.

10. Explain the definition of Data Attributes and provide an example in business.

11. Compare the four measurement levels of data quality attribute by illustrating when each of the four measurement levels should be used to measure the data quality attribute.

12. Explain the definition of Metadata and provide an example in business.

ACTIVITY 2

1. Go to http://data.un.org.

2. Pick **TWO** Asian countries, download the country data, and put the data into **ONE** excel file.

3. Download and install *Tableau Desktop* in your laptop.

   For Mac users, please go to the link below to download the Tableau Desktop: http://downloads.tableausoftware.com/tssoftware/TableauDesktop.dmg

   For Windows users, below is the link for free downloading of the Tableau Desktop:
   http://www.tableausoftware.com/products/desktop/download?os=windows

4. For Mac users, please go to the link below and study the quick user guide on how to connect the Tableau Desktop to Data:
   http://downloads.tableausoftware.com/quickstart/feature-guides/mac.pdf

   For Windows users, please go to the link below and study the quick user guide on how to connect the Tableau Desktop to Data:
   http://downloads.tableausoftware.com/quickstart/feature-guides/connect_data.pdf

5. If you do not have the Microsoft Office System Driver installed in your laptop, please go to the link below to download the driver and install it in your laptop:
   http://www.microsoft.com/en-us/download/confirmation.aspx?id=23734

6. Connect **BOTH** countries' data in the Excel file into the Tableau Desktop in your laptop.

7. Explore and explain what is the difference among connect live, import all data, and import some data options in the data connection process.

8. If we choose the "import some data" option when we connect the data to the Tableau Desktop, how do we change the data filter criteria?

References

Alvarado, A. (2013). *A Very Brief History of Magic [Electronic Illustration]*. Retrieved 17 May 2014 from http://anthonyalvaradoanthonyalvarado.wordpress.com/tag/cave-art.

Barthes, R. (1983). *Empire of Signs*. New York, NY, USA: Hill and Wang.

Belousov-Zhabotinsky. (2014). *Concentric circles caused by a chemical reaction [Electronic Illustration].* Retrieved 24 May 2014 from http://wiki.scienceamusante.net/index.php?title=Belousov-Zhabotinsky.

Bertin, J. (1977). *La graphique et le traitement graphique de l'information*. Flammarion. Paris, France.

Bloom. (2014). *Planetary [Electronic Illustration]*. Retrieved 24 May 2014 from http://planetary.bloom.io.

Bloomberg. (2014). *U.S. Index Metadata [Electronic Illustration]*. Retrieved 24 May 2014 from http://www.bloomberg.com/markets/stocks/world-indexes

Creative Commons Attribution-ShareAlike. (2014). *Contour map of Bryce Canyon National Park, Utah, the U.S.A [Electronic Illustration]*. Retrieved 24 May 2014 from http://en.wikibooks.org/wiki/High_School_Earth_Science/Topographic_Maps

Few, S. (2015). *Signal: Understanding What Matters in a World of Noise.* Burlingame, CA, USA: Analytics Press.

Freeman, R. (2005). *Mixed Reality Toolkit [Electronic Illustration]*. Retrieved 24 May 2014 from http://www0.cs.ucl.ac.uk/staff/rfreeman/demos/demos.htm.

Google Map on SIM University. (2014). *Google Map [Electronic Illustration]*. Retrieved 17 May 2014 from https://www.google.com.sg/maps/place/SIM+University.

Google News. (2014). *Google News [Electronic Illustration]*. Retrieved 17 May 2014 from http://news.google.com.

Lacan, J. (1964). *The Four Fundamental Concepts of Psychoanalysis*. New York, NY, USA: W.W. Norton and Company.

Lévi-Strauss, C. (1963a). *Structural anthropology*. New York, NY, USA: Basic Books.

Marr, D. (1982). *Vision.* New York, NY, USA: W. H. Freeman and Company.

Michell, J. (1997). *Quantitative science and the definition of measurement in psychology.* British Journal of Psychology, 88, 355–383.

Michell, J. (1999). *Measurement in Psychology – A critical history of a methodological concept.* Cambridge, UK: Cambridge University Press.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action.* Oxford Psychology Series 27. Oxford, UK: Oxford University Press.

Murray, G. D. (2016). *Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software. (2nd Edition)* Indianapolis, IN, USA: John Wiley & Sons Inc.

Nelson, G. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN, USA: Hackett Publishing.

Pandey, A. (2013). *Black and White Dots Illusion [Electronic Illustration]*. Retrieved 31 May 2014 from http://www.factrange.com/optical-illusions-will-blow-mind.

QQStay. (2015). *QQStay [Electronic Illustration].* Retrieved 17 May 2014 from http://qqstay.com/.

Saussure, F. de (1959). *Course in General Linguistics.* Reprinted by Fontana/Collins, New York. (Published posthumously based on lectures originally given at the University of Geneva between 1906 and 1911.)

Stevens, S. S. (1946). On the Theory of Scales of Measurement, *103*, 677-680.

Tomonaga, M. (1998). *Tilted Ts and Upright Ts [Electronic Illustration]*. Retrieved 31 May 2014 from http://langint.pri.kyoto-u.ac.jp/ai/intra_data/MasakiTomonaga/Visual_texture_segregation_by_the_chimpanzee.pdf.

Triesman, A. (1980). Preattentive Processing in Vision. *Computer Visions, Graphics and Image Processing, 31*, 156-177.

Tweedie, L. (1997). Characterizing Interactive Externalizations. *CHI'97, ACM, Atlanta Proceedings*, 375-382.

University of New Brunswick. (2014). *3-D Data Visualisation of Colour-Coded Bathymetry Data from Passamaquoddy Bay [Electronic Illustration]*. Retrieved 17 May 2014 from http://www2.unb.ca/gge/News/1996/Aquaculture/PassamaquoddyBay.html.

Ware, C. (2013). *Information Visualization: Perception for Design (Interactive Technologies)* (3rd ed.). Waltham, MA, USA: Elsevier Inc.

Weskamp, M. (2014). *Newsmap [Electronic Illustration]*. Retrieved 17 May 2014 from http://newsmap.jp.

Yau, N. (2013). *Data Points: Visualization that Means Something.* Indianapolis, IN, USA: John Wiley & Sons Inc.

STUDY UNIT 3
FOUNDATION OF THE ART OF DATA VISUALISATION

Learning Outcomes

At the end of this unit, you are expected to:

- compare the eight components of Visual Cues
- explain Trichromatic Theory
- compare CIE, RGB and HSV colour systems
- illustrate the applications of Visual Cues in Data Visualisation
- compare Cartesian, Polar and Geographic Coordinate Systems
- illustrate the applications of Coordinate Systems in Data Visualisation
- compare Linear, Algorithmic, Categorical, Ordinal, Percent and Time Scales
- illustrate the applications of Scales in Data Visualisation
- illustrate the applications of Context in Data Visualisation
- explain Focus-context problem
- compare Distortion, Rapid Zooming, Elision and Multi-Windows techniques to solve Focus-context problem

Overview

This unit introduces the four components of Data Visualisation. In addition, we will also discuss about the applications of the four components in Data Visualisation.

1. Visual Cues

Visual Cues
(Access video via iStudyGuide)

According to Yau (2013), Data Visualisation is built upon data and the four components, which are *Visual Cues, Coordinate Systems, Scales,* and *Context* (refer to Figure 3.1). Sometimes, these four components are explicitly displayed, but other times they may form an invisible framework. These components work together and each of them affects each other.

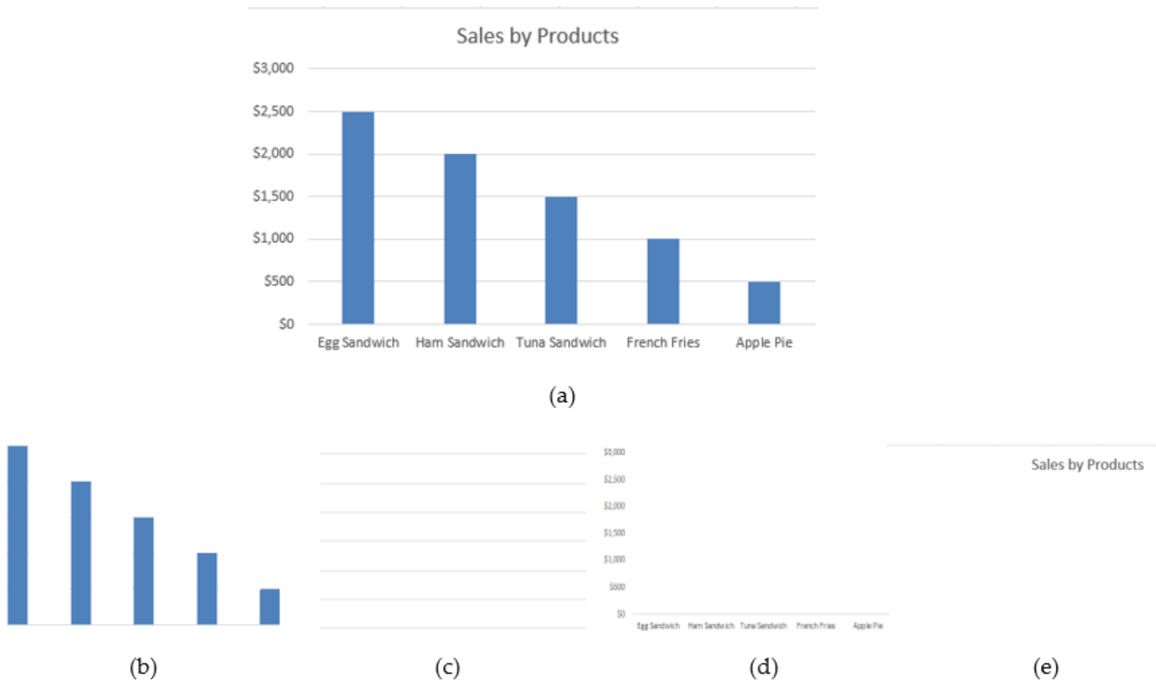


Figure 3.1: (a) Data Visualisation of Sales by products; (b) Visual Cue of the Visualisation; (c) Coordinate System of the Visualisation; (d) Scale of the Visualisation; (e) Context of the Visualisation

In its most basic form, Data Visualisation is simply mapping data onto geometry and colour. It works because the human brain is wired to find patterns and it can switch back and forth between the visual and the number it represents.

We must make sure that the essence of data is not lost in the back and forth between visual and the number it represents. If we cannot map back to the data, the Data Visualisation is just a bunch of shapes. Which Visual Cues we choose depends on data and business objectives. We must choose the right Visual Cue, which changes by purpose, and we must use it correctly. That depends on how we perceive the varied types of Visual Cue, like position, length, angle, etc.

1.1. Position

Position defines where in space or coordinate system the data is. When we use position as Visual Cue, we compare a value based on where the other values are placed in a given space or coordinate system. For example, in Figure 3.2, we use a scatter chart to compare a data point based on its x- and y-coordinates relative to other data points.

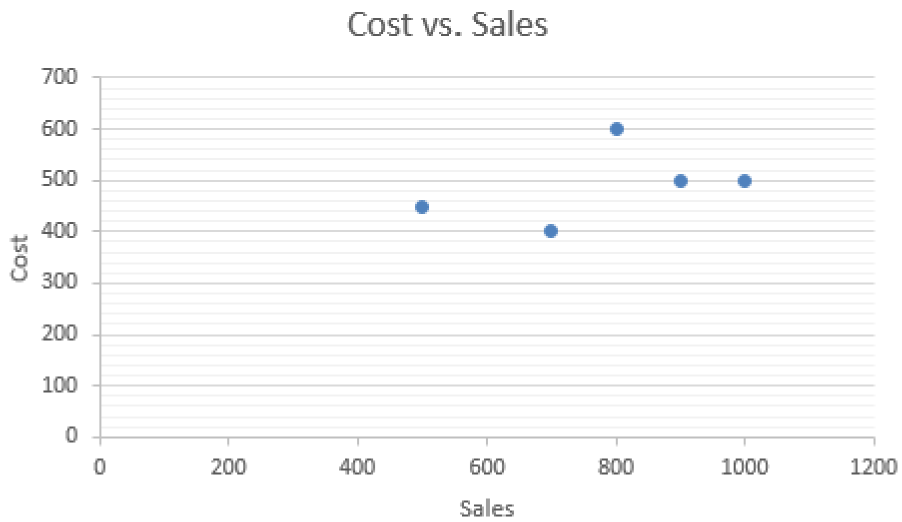


Figure 3.2: Cost versus sales shown using Scatter Chart, where Position is used as Visual Cue

The above scatter chart uses position as a Visual Cue so that we can compare the cost of the product when the sales equals to $500, and the cost of product when the sales equals to other values, like $700, $800, etc.

Position Visual Cue tends to take up less space than the other types of Visual Cue, because we can draw all the data points within the x and y planes, and represent each data point using a symbol like dot, cross, square or diamond.

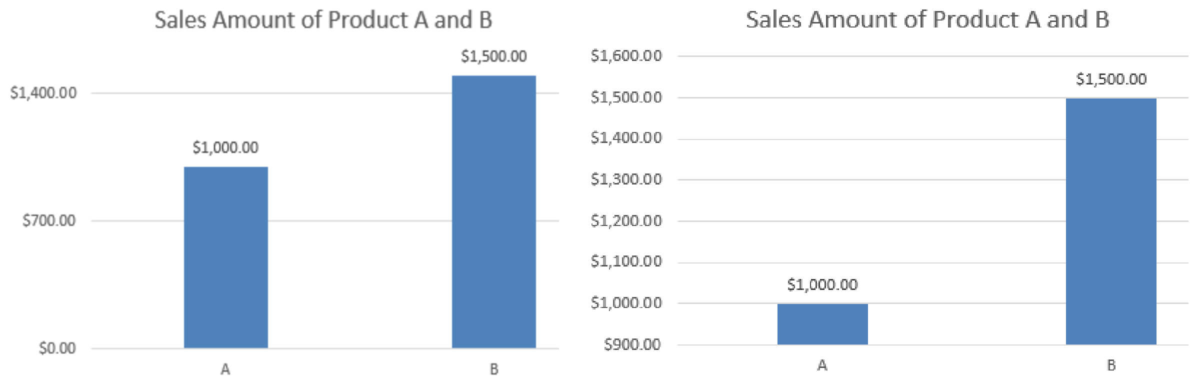
Unlike other types of visual cue that take up space to compare values, all data points in position-based Data Visualisation take up the same amount of space (i.e. symbol size). By using Position Visual Cue, we can spot trends, clusters and outliers by plotting a lot of data points within the x and y planes. However, it will be challenging to identify what each data point represents. Even in the interactive scatter plot, we still need to mouse over or select a data point to get more information. The overlapping of two or more data points can cause more challenge in identifying what the data points represent.

1.2. Length

Length is the most commonly used Visual Cue to compare data values in several chart types, like bar chart, bullet chart, combo chart, etc. The longer a bar or a bullet point, the greater the absolute value. It can work in both horizontal and vertical directions.

To judge the length visually, we measure the distance from one end of a shape to the other end. In this case, to compare data values based on lengths, we must see both ends of the bars or bulletspoints. Otherwise, we may end up having a skewed view of maximums, minimums, and everything in between.

For example, as shown in Figure 3.3, we use two bar charts with Length Visual Cue to display the same sales amount of two products of a company.



(a) (b)

Figure 3.3: Sales amount of products A and B shown using Length Visual Cue

If both bar charts are shown quickly, the difference in sales amounts between Product A and Product B in Chart (b) will be perceived to be greater than that of Chart (a). This is because on Chart (a), the axis starts at zero and its major unit is labelled every $700. Whereas in Chart (b), the axis starts at $900 and its major unit is labelled every $100.

1.3. Angle

Angle, as a Visual Cue, ranges from zero to 360 degrees on a circle. A straight line is 180 degrees, an obtuse angle is an angle between 90 degrees and 180 degrees, and an acute angle is an angle of less than 90 degrees.

For each angle in between zero and 360 degrees, there is an implied opposite angle that completes the rotation. The opposite angle is called the conjugate of an angle.

This is why angles are commonly used to represent parts of a whole, as shown in Figure 3.4.

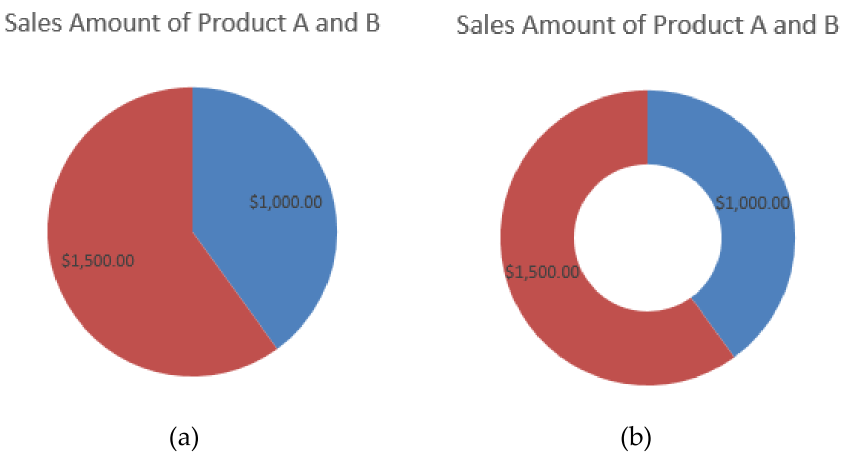


Figure 3.4: Sales amount of products A and B shown using Angle Visual Cue

In Figure 3.4, Chart (a) is a pie chart that shows the proportion of sales amounts of Product A and Product B, where the relative degrees of $1,000.00 (i.e. Product A's sales amount) and $1,500.00 (i.e. Product B's sales amount) in the circle is the Angle Visual Cue. Chart (b) is a doughnut chart that shows the proportion of sales amounts of Product A and Product B. The length of the arc is the Angle Visual Cue because the centre of the circle, which indicates angles, is removed.

1.4. Direction

Direction as a Visual Cue, is similar to angle. Instead of relying on two vectors joined at a point, direction relies on a single vector's orientation in a coordinate system to see which way is up, down, left, right, and everything in between. Direction helps us determine the slope to see increases, decreases and fluctuations as shown in Figure 3.5.

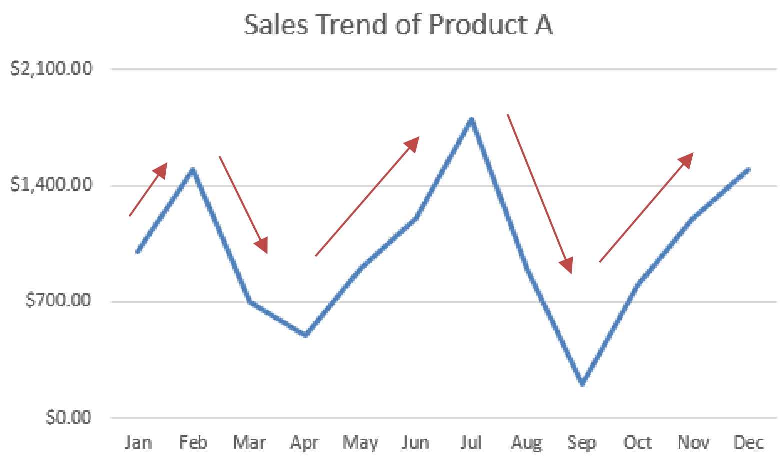
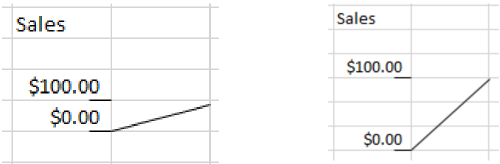


Figure 3.5: Sales trend of Product A shown using line chart, where Direction is used as Visual Cue

The amount of perceived changes depends a lot on the scale, as shown in Figure 3.6. For example, we can make a small change in percentage looks like a lot by stretching out of scale. Likewise, we can make a big change looks like a little by compressing the scale.



(a) (b)

Figure 3.6: The same amount of change shown using different scales

The rule of thumb is to scale the visualisation so that the direction fluctuates mostly around 45 degrees, but this rule is not cast in stone. We can start with this rule and adjust accordingly base on the context. If a small change is significant, then it may be appropriate to stretch the scale so that we can see the shift. In contrast, if a small change is insignificant, we should not stretch out of scale so that the change will not look major.

1.5. Shape

Shape or symbol is commonly used as a Visual Cue to differentiate categories and objects. When we use shapes on a map, they can represent things in the real world. For example, we may use trees as a symbol to represent forests, or use houses as a symbol to represent residential areas. In a chart, we may use triangles and squares in a scatter chart to differentiate different categories of data, as shown in Figure 3.7. Nevertheless, various shapes can provide a context that data points alone cannot.

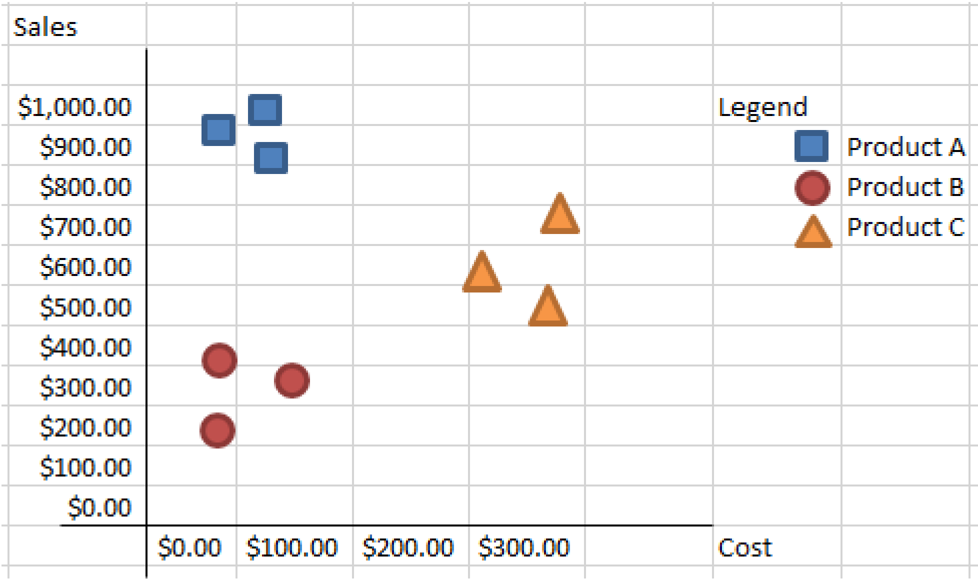


Figure 3.7: Different shapes represent different product on a Scatter Chart

1.6. Area and Volume

When we use *Area* and *Volume* as Visual Cues, bigger objects represent greater values. For instance, Length, Area, and Volume can be used to represent data with sizes (i.e. amount of space) in either two or three dimensional space. When we use circles (i.e. two dimensional spaces) or spheres (i.e. three dimensional space) as Shape Visual Cues, we can also use different sizes of circle or sphere to represent different data values.

The most common mistake is to scope a two or three dimensional object using one of the dimensions, such as height, without keeping the proportion of all dimensions. This will result in shapes that are either too big or too small, which makes it impossible to fairly compare the data values.

For example, when we use a square shape with two dimensions (i.e. width and height) as a Visual Cue to represent our data, and we apply Area Visual Cue, we know that the greater the value, the greater the area of a square. If one value is 50 percent greater than another, we want the area of the square to be 50 percent greater than the other. However, if we increase the width and height of the square by 50 percent, instead of keeping the proportion of width and height, the larger square are will be increased by 125 percent, as shown in Figure 3.8.

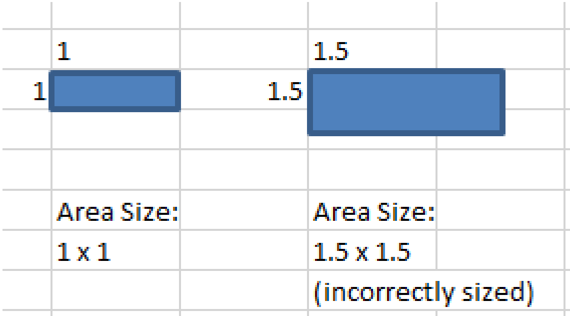


Figure 3.8: Squares sized by different dimension values

We will run into the same problem with three dimensional objects, but the mistake will be more pronounced. For example, when we increase the width, height and depth of a cube by 50 percent, the volume of the larger cube will be increased by approximately 238 percent.

1.7. Colour

Colour, as a Visual Cue helps us break camouflage because things differ visually from their surroundings only by their colours. For example, if we have a colour vision, we can easily see that there is a black bear in Figure 3.9. If we do not have a colour vision, it will be much harder for us to spot the black bear. Clearly, this can be a life-or-death decision for the hunters or for a person who walks in the jungle.

Figure 3.9: The importance of having Colour as Visual Cue (Sears, 2014)

The role that colour plays ecologically suggests ways that it can be used in data visualisation. Colour is good for labelling and categorisation, but poor for displaying shape, detail or space.

1.7.1. Trichromacy Theory

According to Young (1802) and Helmholtz (1866), human beings have three distinct colour receptors, called *Cones*, in our retinas that are active at normal light level. We also have *Rods* that are sensitive at low light level, and overstimulated in all but the dimmest light where it influences can be ignored. Thus, in order to understand colour vision, we need to consider only the cones. The three distinct colour receptors or cones is the reason for the basic three-dimensionality of human colour vision, called *Trichomacy*.

Figure 3.10 shows the cone sensitivity functions. The plots show how light of different wavelengths is absorbed by the different receptors. It is clear that two of the functions, which peak at 540 nanometers and 580 nanometers respectively, overlap considerably. The third wavelength is much more distinct, with peak sensitivity at 450 nanometers.

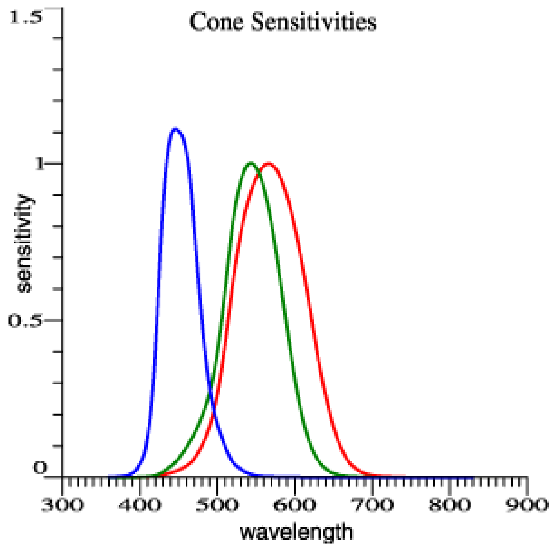


Figure 3.10: Cones Sensitivity Functions (Worthey, 2010)

As we only have three different receptor types involved in a colour vision, it is possible to match a particular patch of coloured light with a mixture of just three colours, called *Primary Colours*. It does not matter that the target patch may have a completely different spectral composition. The only thing that matters is that the matching primary colours are balanced enough to produce the same response from the receptors as the patch of light to be matched.

1.7.2. Colour Blindness

About 10% of the male population and about 1% of the female population suffer from colour vision deficiency. Most colour deficiencies are caused by the lack of either long-wavelength-sensitive cones called *Protanopia*, or the medium-wavelength-sensitive cones called *Deuteranopia*.

Both Protanopia and Deuteranopia result in the inability to distinguish red and green. One way of describing colour vision deficiency is by collapsing the three-dimensional colour space of normal colour vision into a two-dimensional colour space, as shown in Figure 3.11.

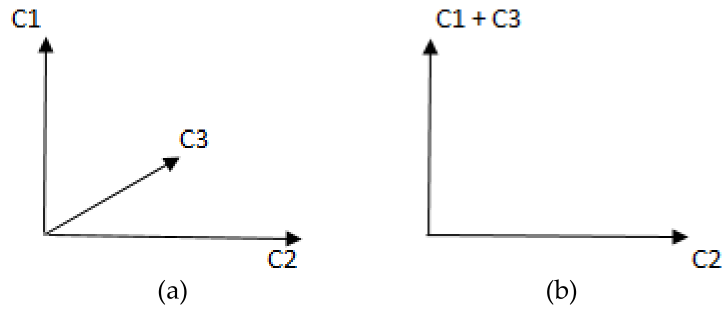


Figure 3.11: (a) Cone response space, defined by the response of each of the three cone types; (b) The space becomes two-dimensional in the case of the common colour vision deficiencies

If colour is widely used for Data Visualisation, we immediately create a new class of people with a disability. Colour blindness already disqualifies applicants for jobs, such as pilots, because of the need to distinguish colour-coded lights.

1.7.3. Colour Systems

According to Stockman (2003), *Colourimetry* is the measurement of the physical intensity of colours as opposed to their subjective brightness. The fact that we can match any colour with a mixture of no more than three lights, usually called primary colours, is the basis of Colourimetry. *Colour Model* is an abstract mathematical model that describes the ways colours can be represented as tuples of numbers.

We can describe a colour by the following colour model:

$$C \equiv rR + gG + bB$$

Where

C is the colour to be matched

R, G and B are the primary sources to be used to create a match

r, g and b are the amounts of each primary sources

≡ denotes a perceptual match, which means the sample and the mixture of red, green and blue (primary colours) look identical

Supposed we have a coloured light that cannot be matched, because it is outside the gamut of three primary sources. We can still achieve a match by adding part of one of the primary sources to the sample source. For example, if the red projector is redirected to the sample source, we will have

$$C + rR \equiv gG + bB$$

which can also be written as

$$C \equiv -rR + gG + bB$$

Once we allow the concept of negative values for the primary sources, it becomes possible to state that any coloured light can be matched by a weighted sum of any three distinct primary colours.

This concept is called RGB colour system. It is illustrated in Figure 3.12. In the colour-matching experiment, three projectors are completely overlapped. The observer should be able to match the overlapped colour with the colour generated from the test projector. For example, to match lilac from the test projector, the red, green and blue projectors are adjusted so that a large amount of light comes from the red and blue projectors, with only a small amount of light from the green projector.

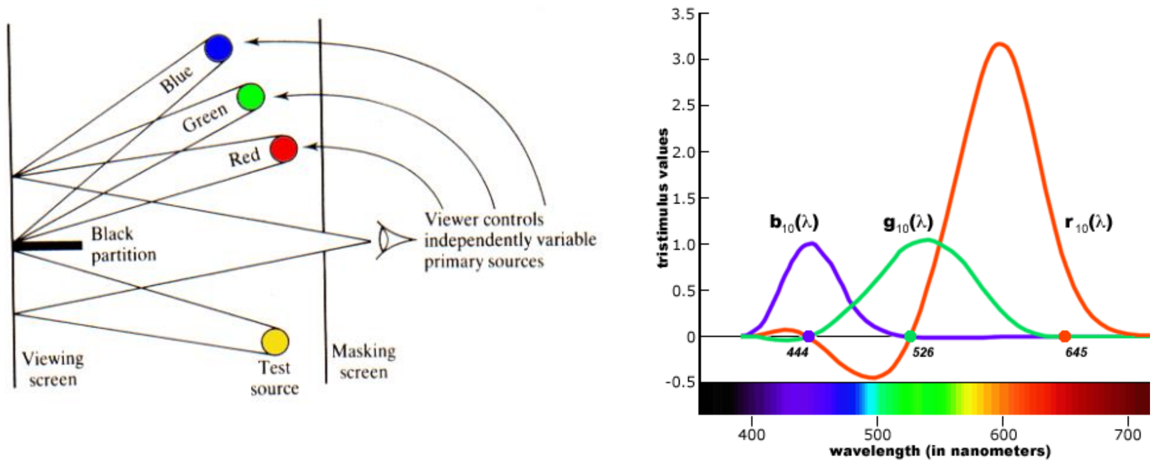


Figure 3.12: (a) Cone response space, defined by the response of each of the three cone types; (b) The space becomes two-dimensional in the case of the common colour vision deficiencies (MacEvoy, 2005)

With the RGB colour system, to give someone a precise colour specification with the standard primary colours, we simply need to make a match, then send that person the amounts of each of the three primary colours needed for the match. The person can then adjust the standard lamps to reproduce the colours. This approach is theoretically sound, but it is not very practical to create standard primary lamps, because it will be very difficult to maintain and calibrate the lamps.

To solve this issue, we can define a hypothetical person whose colour sensitivity functions are held to be typical of all humans. We assume that everyone has the same receptor functions, and we create a set of abstract primary lamps based on the human receptor characteristics. The Commission Internationale de L'Eclairage (CIE) colour system was made prior to 1931 based on this concept.

The CIE colour system uses a set of abstract primaries called Tristimulus values that are labeled XYZ, as shown in Figure 3.13. These values are chosen for their mathematical properties, and not because they match any set of actual lights. One important feature of the CIE colour system is that the Y Tristimulus value is the same as the luminance.

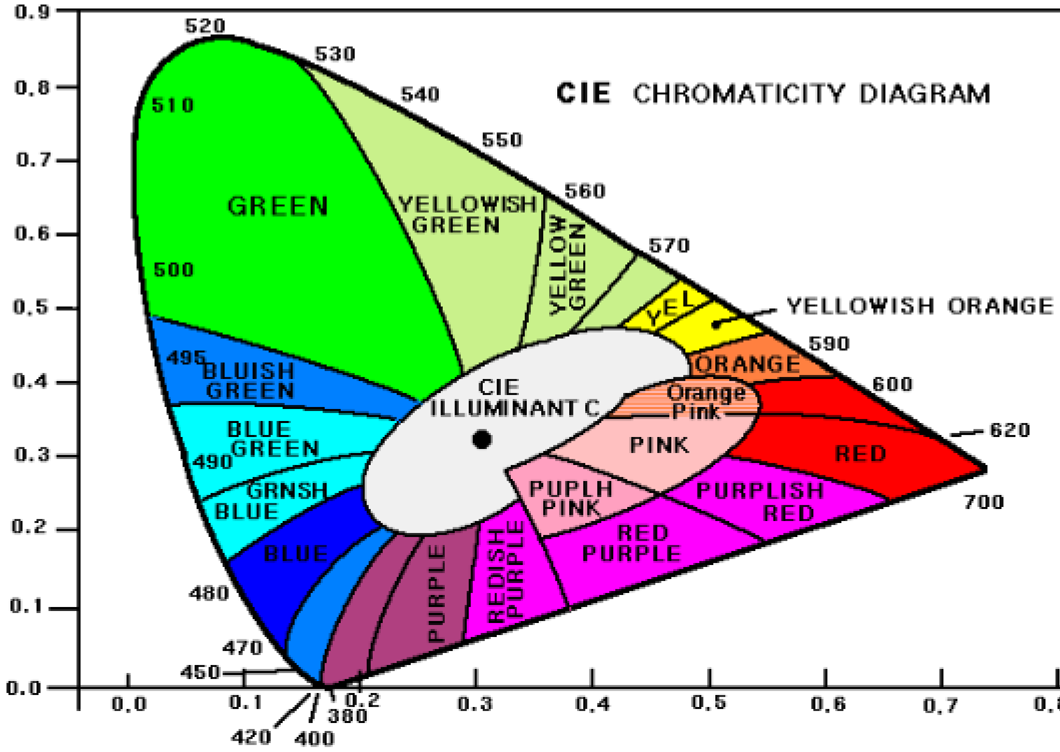


Figure 3.13: CIE Chromaticity Diagram (Glen, 2012)

The CIE colour system is by far the most widely adopted colour system to measure coloured lights. We should always use the CIE colour system when precise colour specification is required. For example, if we would like to reproduce a colour generated on one Liquid Crystal Display (LCD) monitor on the second LCD monitor, the best procedure would be first to convert the colours into the CIE Tristimulus values, then convert the Tristimulus values into the RGB colour space of the second LCD monitor.

It is more difficult to specify the surface colours than to specify the lights, because we need to take into account an illuminant. Unlike lights, colour pigments are not additive. Therefore, another type of colour system which is known as HSV colour system emerged.

The HSV colour system uses colour hue, colour saturation, and black-white brightness to specify the surface colours. This colour system is often used by artists. It is often more natural for them to think about a colour in terms of hue and saturation than in terms of primary, additive or subtractive colour space or CIE TriStimulus values. The HSV colour system, as shown in Figure 3.14, is a transformation of an RGB colour space, where its components and Colourimetry are relative to the RGB colour space from which the HSV colour space was derived.

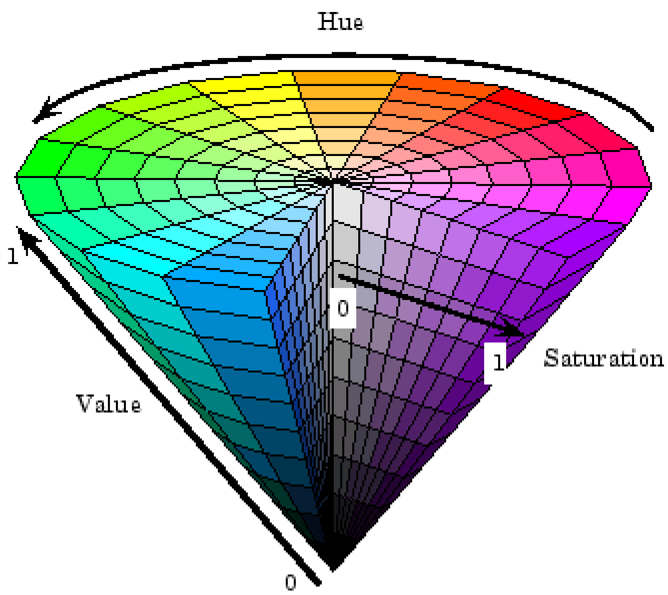


Figure 3.14: HSV Colour Space (The MathWorks Inc., 2014)

In the HSV colour system, *Hue* refers to which part of the rainbow colour map a colour belongs to, such as red or green. *Saturation* refers to how rich a Colour Hue is, for example, neon colours are very saturated, while pastel colours are less saturated. Value denotes how bright a colour is, or in other words, how close a colour is to black or white.

In Data Visualisation, it is very common to encode value as a hue shifts from blue (i.e. cold or low) to red (i.e. hot or high). The resulting colour map is usually referred to as the rainbow colour map as shown in Figure 3.15.

Figure 3.15: Rainbow Colour Map (Stam, J., 2007)

However, research shows that the rainbow colour map is a poor choice in most Data Visualisation tasks, as it tends to hide variations in regions of low contrast and appears less smooth in regions of high contrast, as shown in Figure 3.16.

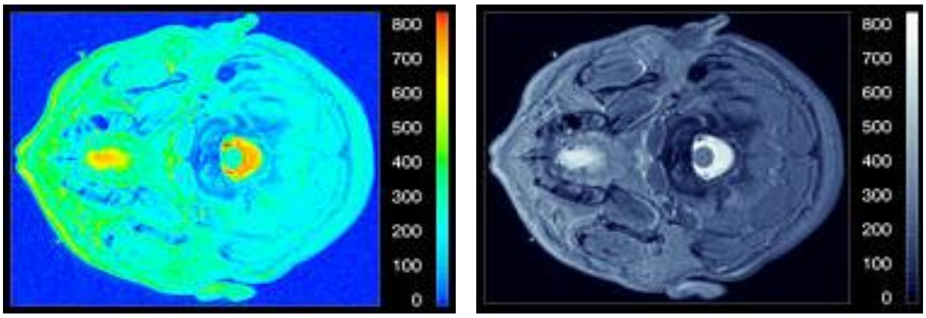


Figure 3.16: The rainbow colour map (left) tends to hide variation in regions of low contrast and should be used with care. One option is to use an isomorphic colour map (right) where equal steps in the data correspond to equal perceptual steps (Rogowitz & Treinish, 1996)

Unfortunately, there isn't any colour map that can be recommended for all Data Visualisation situations. It is, therefore, important to choose a suitable colour map so that it corresponds to the analysis or task at hand.

2. Coordinate Systems

WATCH

Coordinate Systems
(Access video via iStudyGuide)

According to Yau (2013), when we create Data Visualisation, we need to place the objects somewhere. *Coordinate System* is a Data Visualisation component that determines the position of the objects (e.g. data points, shapes or symbols). There are three most basic coordinate systems: *Cartesian*, *Polar* and *Geographic* coordinate systems.

2.1. Cartesian

Cartesian coordinate system, as shown in Figure 3.17, is a coordinate system that specifies each data point on a plane by a pair of numerical coordinates. The numerical coordinates are the signed distances from the data point to the two fixed perpendicular reference lines, which are measured in the same unit of length. These reference lines are called x- axis and y-axis. Both axes meet at a point, called the *origin*, which is usually represented by ordered pair (0, 0). The numerical coordinates can also be expressed as a signed distance from the origin.

We can also use the same concept to define the position of any data point in *n*-dimensional space by *n* Cartesian coordinates. The coordinates are the signed distance from the data point to the *n* fixed perpendicular reference lines which are measured in the same unit of length.

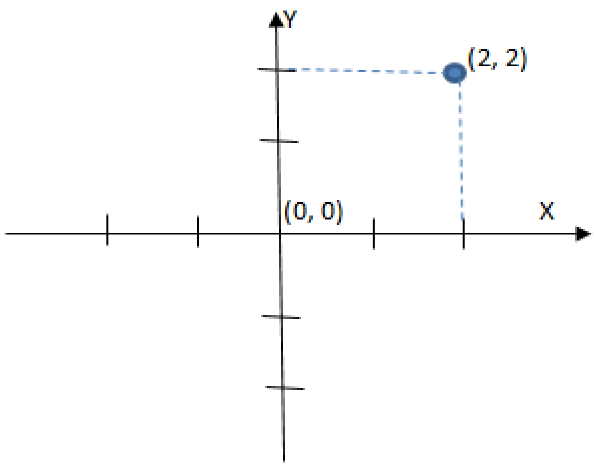


Figure 3.17: Cartesian coordinate system

2.2. Polar

Polar coordinate system, as shown in Figure 3.18, is a two-dimensional coordinate system. Each data point is determined by the distance between a fixed point and an angle from a fixed direction. The fixed point, which is analogous to the origin in the Cartesian coordinate system, is called the *Pole*. The ray or half-line from the Pole in the fixed direction is called the *Polar Axis*. The distance from the Pole is called the *Radial Coordinate* or *Radius*, and the angle is called the *Angular Coordinate*, *Polar Angle* or *Azimuth*.

This coordinate system is less used than the Cartesian coordinate system, but it is useful in cases where the angle or direction is important.

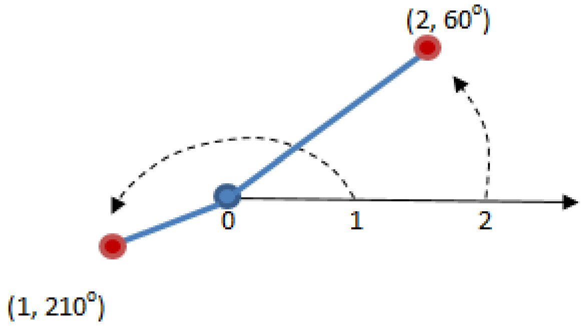


Figure 3.18: Polar coordinate system

2.3. Geographic

Geographic coordinate system, as shown in Figure 3.19, is a coordinate system that enables every location on the earth to be specified by a set of numbers or letters. To represent the location, the coordinate system commonly uses latitude and longitude -- angles relative to the Equator and the Prime Meridian, respectively. Sometime the coordinate system may also use elevation.

Latitude lines run east and west, which indicates north and south positions on the globe. *Longitude lines* run north and south, which indicates east and west positions. *Elevation* can be thought of as a third dimension. Compared with Cartesian Coordinate System, latitude is like the horizontal axis, and longitude is like the vertical axis.

The surface of the earth is wrapped around a spherical mass, but we usually want to display a location on earth on a two-dimensional surface, like a piece of paper or a computer screen. Therefore, there is a variety of ways to map the surface of the Earth on a two-dimensional surface, which are called *projections*.

2.3.1. Equirectangular

It is typically used for thematic mapping and it does not preserve any area or angel.

Figure 3.19: Equirectangular mapping way (TUBS, 2011)

2.3.2. Albers

It does not preserve scale and shape, and angle is minimally distorted.

Figure 3.20: Albers mapping way (Knippers, 2009)

2.3.3. Mercator

It preserves angles and shapes in small area, so it is good for direction.

Figure 3.21: Mercator mapping way (Waterman, 1999)

2.3.4. Lambert Conformal Conic

It is better used for showing smaller areas and it is often used for aeronautical maps.

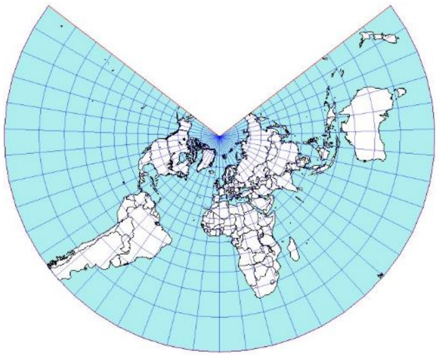


Figure 3.22: Lambert Conformal Conic mapping way (Aquarius.net, 2002)

2.3.5. Sinusoidal

It preserves area and it is useful for showing areas near the prime meridian.

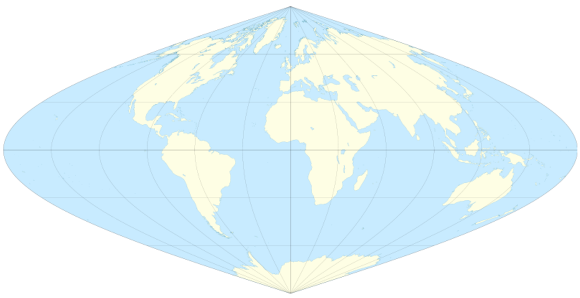


Figure 3.23: Sinusoidal mapping way (Gaba, 2008)

2.3.6. Polyconic

It was used to show the map of the U.S. in the mid-1900s. There are little distortions in small areas near the meridian.

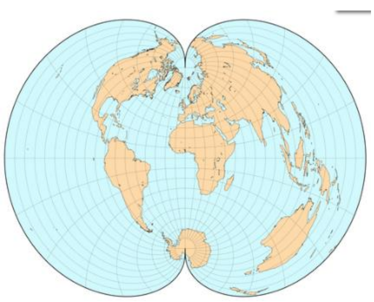


Figure 3.24: Polyconic mapping way (Furuti2013)

2.3.7. Winkel Tripel

It is a good choice for showing the world map, because it minimises area, angle and distance distortions.

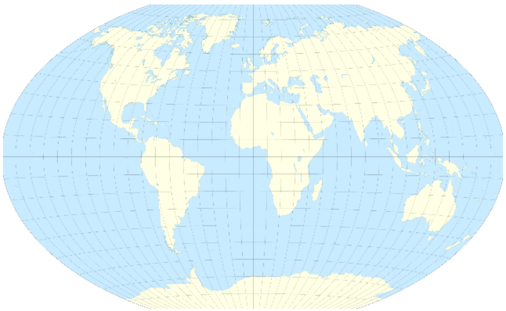


Figure 3.25: Winkel Tripel mapping way (Mapsof.net, 2014)

2.3.8. Robinson

It is a good choice for showing the world map because it compromises preserving areas and angles.

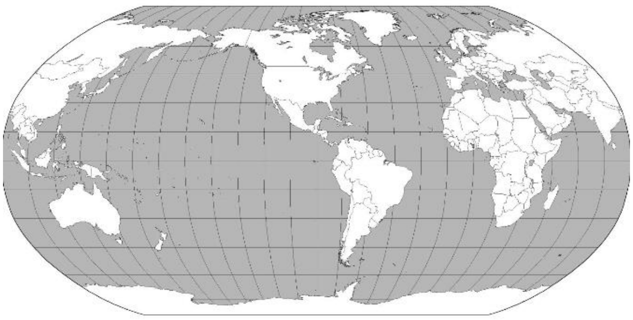


Figure 3.26: Robinson mapping way (Valentine, 2013)

2.3.9. Orthographic

It represents a three-dimensional object in a two-dimensional space. Using this method, the user needs to rotate to the area or to the location of interest.

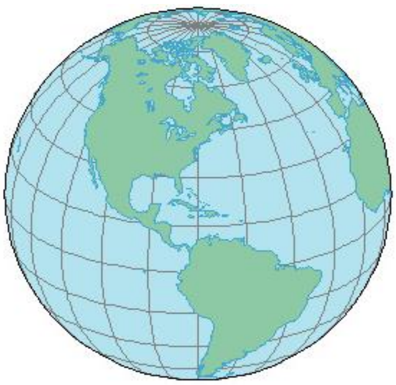


Figure 3.27: Orthographic mapping way (Krygier, 2014)

3. Scales

Scales is a Data Visualisation component that defines where the shapes are placed and how they are shaded.

There are several types of scales as listed below.

3.1. Linear Scale

In *Linear Scale*, the visual spacing between each of the data points is the same regardless where the data points are on the axis.

For example, a Customer Relationship Manager of the apparel company would like to visualise customer satisfaction on the company's products. He can use the Linear Scale (i.e. 0 represents very disappointed and 5 represents very satisfied), as illustrated in Figure 3.28 to visualise the data. If we were to measure the distance between two data points on the lower end of the Linear Scale, it would be the same as the distance between two data points on the higher end of the Linear Scale.

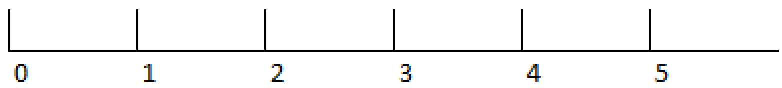


Figure 3.28: Linear Scale

3.2. Logarithmic Scale

Logarithmic scale condenses the distance between each of the data points when the value of the data points increase. This scale is useful if we are interested to visualise the percentage difference or the rate of change amongst each of the data points, especially when the data has a wide range.

For example, the sales amount of a food and beverages company like Burger King in year 2000 was S$ 1,000 and doubled each year.

The following charts show the sales amount on Linear and Logarithmic Scales:

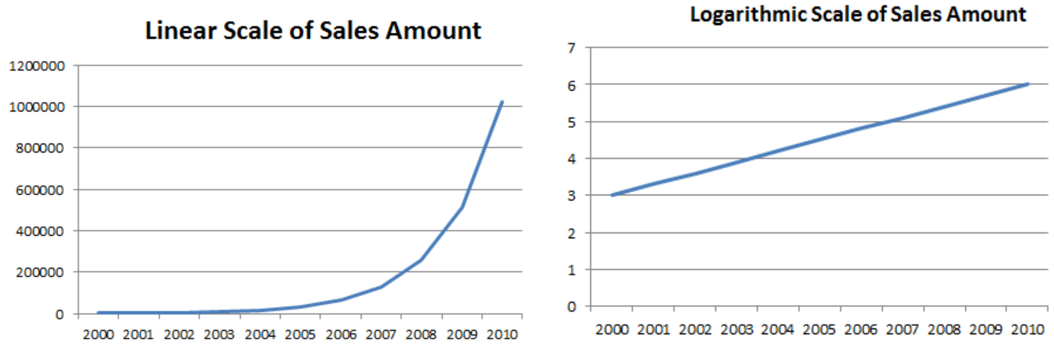


Figure 3.29: Linear Scale versus Logarithmic Scale

The Linear Scale shows the absolute number of sales over time, while the Logarithmic Scale shows the rate of change of the sales amount over time. By visualising the data using the Logarithmic Scale, it is much clearer that the rate of change is constant. Although the Logarithmic Scale is extremely useful, it is not easily understood by all. Thus, the Data Visualisation designer must know their target audience.

3.3. Percent Scale

Percent Scale is usually linear, but when it is used to represent part of the whole data, its maximum is 100 percent. In addition, the sum of percentages should not exceed 100 percent.

For example, using the food and beverages company scenario, the company sales in year 2000 was S$ 1,000 in the Singapore market, S$ 500 in the Malaysia market, and S$ 500 in the Thailand market. The following chart show the sales amount in Percent Scale:

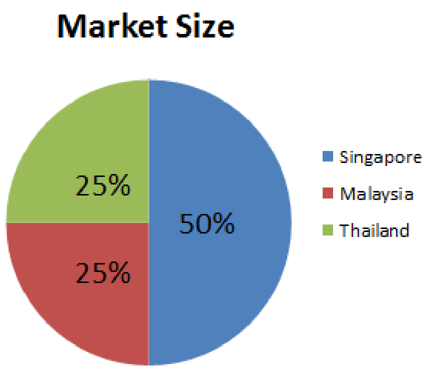


Figure 3.30: Percent Scale

3.4. Categorical Scale and Ordinal Scale

We use *Categorical Scale* when we want to provide visual separation of categorical data, such as country of residence or gender. The Categorical Scale often works with a Linear Scale or a Logarithmic Scale.

For example, a Sales Manager of an apparel company would like to visualise product sales in year 2000. The Sales Manager can use a Bar Chart as shown in Figure 3.31, where it uses a Categorical Scale on the horizontal axis; and a Linear Scale on the vertical axis to show the sales amount for different years.

The spacing between each category (i.e. each bar) is arbitrary because it does not depend on the numeric value, it is usually adjusted to provide clarity.

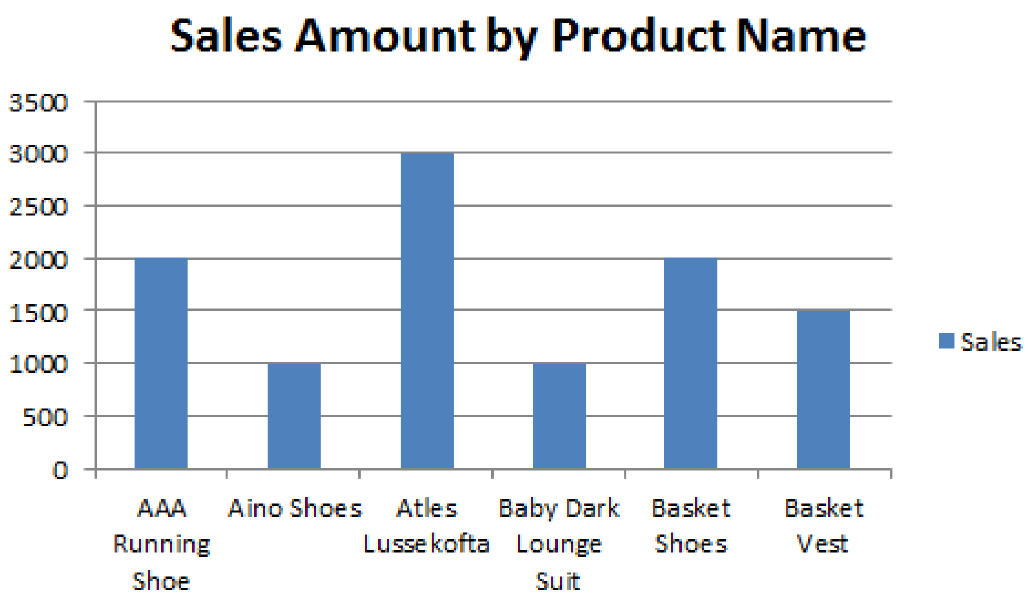


Figure 3.31: Categorical Scale works with Linear Scale

When the order of the categorical data matters, we should order the Categorical Scale in the context of the data from year 2000 to the latest year, which is year 2010 in Figure 3.32. This type of scale is known as an *Ordinal Scale*.

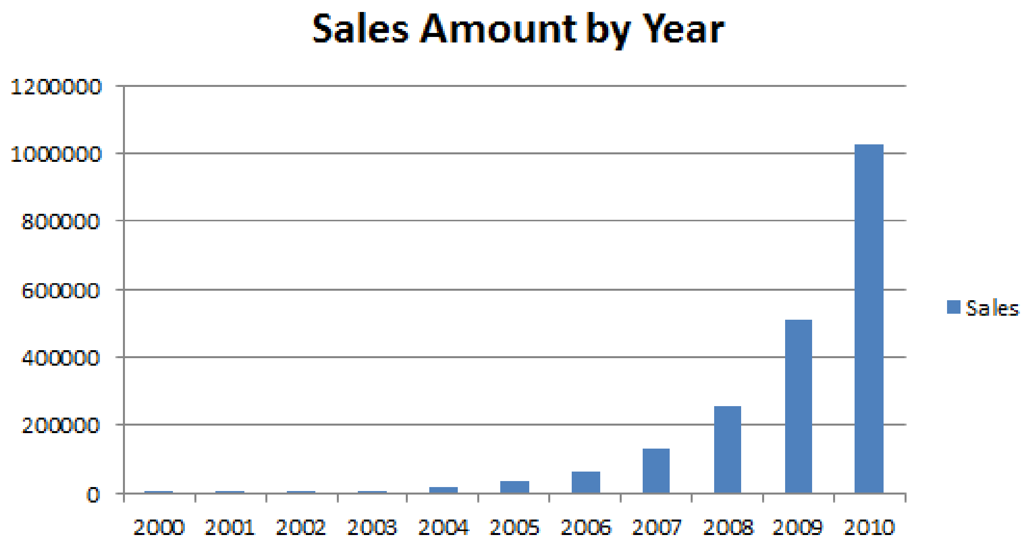


Figure 3.32: Ordinal Scale works with Linear Scale

3.5. Time Scale

We use the *Time Scale* when we want to plot temporal data on a Linear Scale, or to divide the Temporal Data into categorical scale, such as by year, month or day. Figure 3.32 is an example of using Time Scale (i.e. year) in a Bar Chart. When communicating data to the audience, the Time Scale gives us the advantage of lending the reader connection, because time is a part of everyday life.

4. Context

Context is a Data Visualisation component that lends to better understanding of who, what, when, where and why of the data. Context can make the data clearer for interpretations.

When we would like to enable Data Visualisation viewers to see the Data Visualisation object of primary interest in full detail, and at the same time get an overview within the context (i.e. surrounding information) available, this is what we called a *focus-context problem* in data visualisation.

Card, Mackinlay and Shneiderman (1999) explained that focus-context problem starts from three premises: First, the viewer needs both context (i.e. overview of the information), and focus (i.e. detail of the information) simultaneously. Second, the information needed in the overview may be different from that needed in detail. Third, these two types of information need to be combined within a single interactive data visualisation. (p. 1-34)

A focus-context Data Visualisation therefore allows viewers to have the information of interest in the foreground and all the remaining information in the background which are simultaneously visible (i.e. seeing the trees without missing the forest).
The focus-context problem is not always spatial related problem. There are also structural and temporal related problems.

Spatial related problems are common to all Data Visualisation that use maps. For example, a market researcher may wish to understand the behaviour of the individual customer of a shop in Ang Mo Kio. This information is understood in the context of the distance between the shop and the other shops, and the distance between the shop and the customers.

Structural related problem arises when we try to visualise data that have structural components at many levels. Suppose a company sells their products in the U.S. and internationally. We wish to visualise the sales performance of their stores. We may need to understand the company's structure from country, state or province level to city or town level of details.

Temporal related problem involves understanding the timing of data at very different scales. For example, to visualise the performance of Apple shares in the U.S. share market, it can be useful to know the overall performance of the Apple share on a daily basis. However, it can also be useful to analyse the share performance on hourly, weekly, monthly or yearly basis.

We are fortunate that the human brain is able to integrate detailed information from successive fixation of the fovea, with the less detailed information that is available at the periphery of vision. This is combined with the information that comes from the prior sequence of fixations. For each new fixation, our brain will somehow match the key objects in the previous view with those same objects that are moved to new locations. Differing level of details is supported in normal perception because objects are seen at much lower resolution at the periphery of vision than in the fovea. In addition, we have no difficulty in recognising objects at different distances, which means that scale-invariance operations are supported in normal perception. Therefore, solutions to the focus context problem are by taking advantage of these perceptual capabilities.

There are four different visualisation techniques to solve the focus context problem: Distortion, Rapid Zooming, Elision, and Multiple Wdindows.

4.1. Distortion

The *Distortion* technique spatially distorts a data presentation to give more room to the designated points of interest, and to decrease the space given to regions away from those points. What is of specific interest is spatially expanded at the expense of what is not, thus providing both focus and context. Figure 3.33 shows one of such methods called *Hyperbolic Tree Browser* (Lamping, Rao & Pirolli, 1995), where parts of the graph are dynamically repositioned and resized based on selected points of interest. The selected node is expanded to show its content.

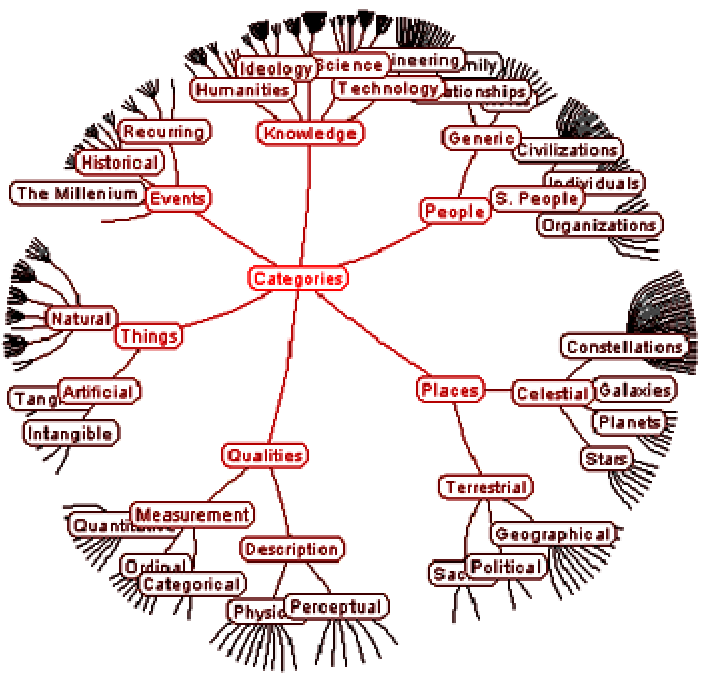


Figure 3.33: The Hyperbolic Tree Browser (Heer, 2004)

Table lens (Rao & Card, 1994) also uses the distortion technique to expand multiple foci simultaneously, as shown in Figure 3.34. It fuses symbolic and graphical representations of data into a single understandable view that can be fluidly adjusted by the viewers.

Figure 3.34: The table Lens (Friendly, 2009)

4.2. Rapid Zooming

The *Rapid Zooming* technique allows viewers to zoom rapidly in and out of points of interest. This technique provides a large information landscape, even though only a part of them is visible in the viewing window at any one time. This means that the focus and context are not simultaneously available, but viewers can move rapidly and smoothly from focus to context and vice versa. With rapid smooth scaling, viewers can perceptually integrate the information over time. Figure 3.35 illustrates the Pad and Pad++ systems (Bederson and Hallan, 1994) which are based on this principle. Such systems provide a large planar data landscape, with an interface of simple point-and-click technique to move rapidly in and out.

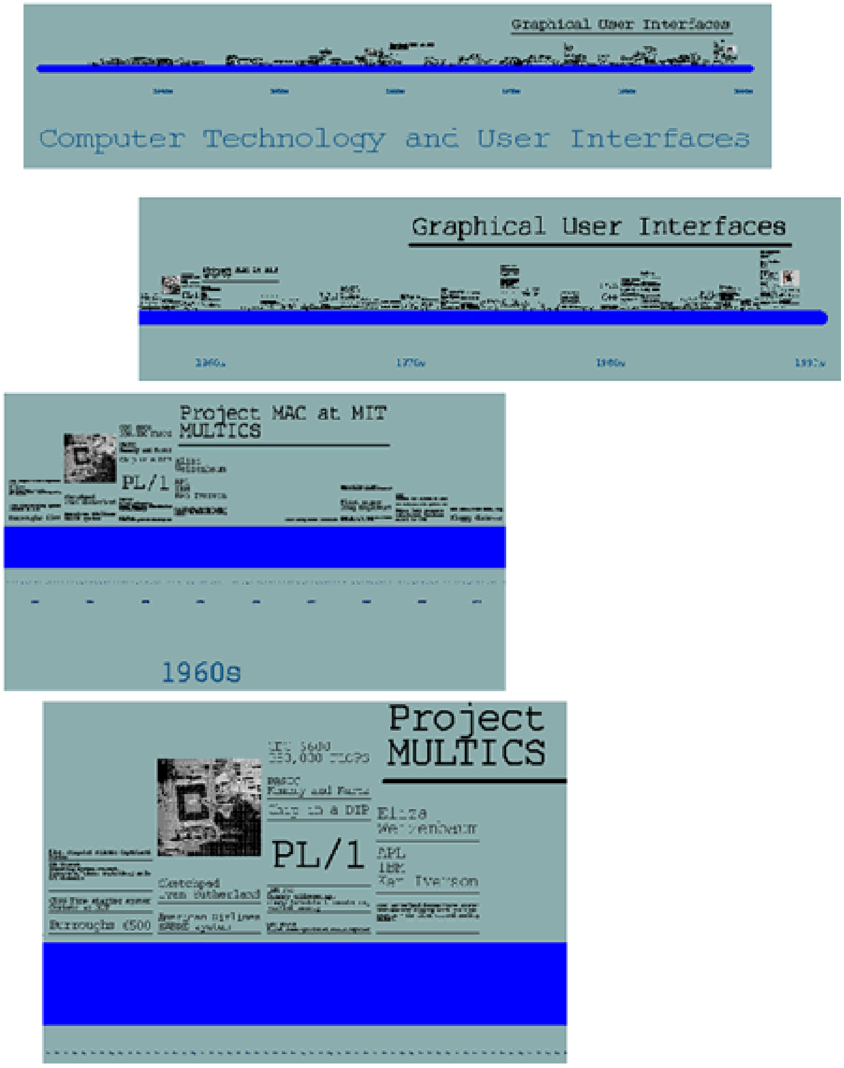


Figure 3.35: Sequence views of Pad++ system (The University of New Mexico, 1996)

4.3. Elision

The *Elision* technique hides parts of a structure from viewers until they are needed. Typically, this can be achieved by collapsing a large graphical structure into a single graphical object. This technique can be applied to text as well as to graphics.
Figure 3.36 illustrates the generalised, fish-eye technique of viewing data (Furnas, 1986), where lesser and lesser detail is shown as the distance from the focus of interest increases.

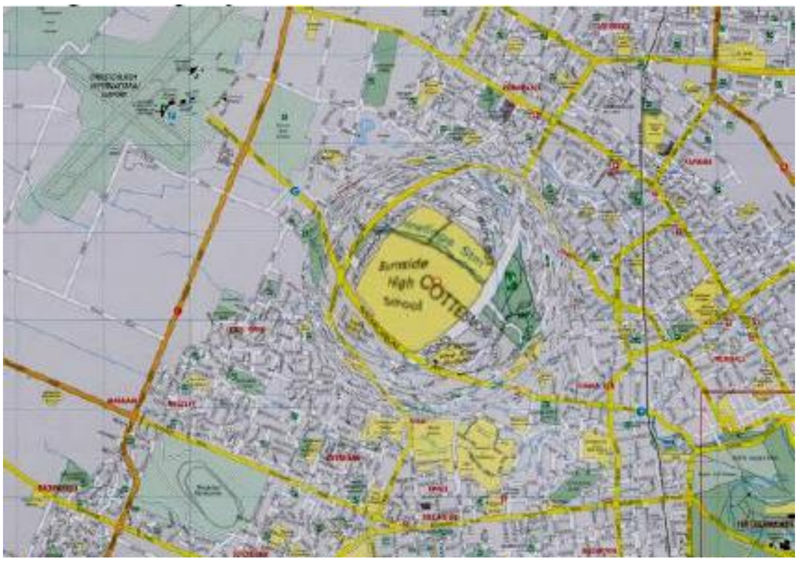


Figure 3.36: The generalised Fish-Eye technique for viewing data (Cockburn, Karlson, & Bederson, 2007)

4.4. Multiple Windows

The *Multiple Windows* technique allows viewers to have one window that shows an overview of the data, and several other windows that show the expanded details. The major perceptual problem with the Multiple Windows technique is that detailed information in one window is disconnected from the overview (i.e. the context) as shown in another window. A solution for this problem is to use lines or transparent overlay to connect the boundaries of the zoomed window to the source window.

Figure 3.37 illustrates the Spiral Calendar (Card, Mackinlay & Shneiderman, 1994, p. 1-34), where information of one calendar (i.e. zoom window or focus) is linked to the source window (i.e. the context) by a connecting transparent overlay.

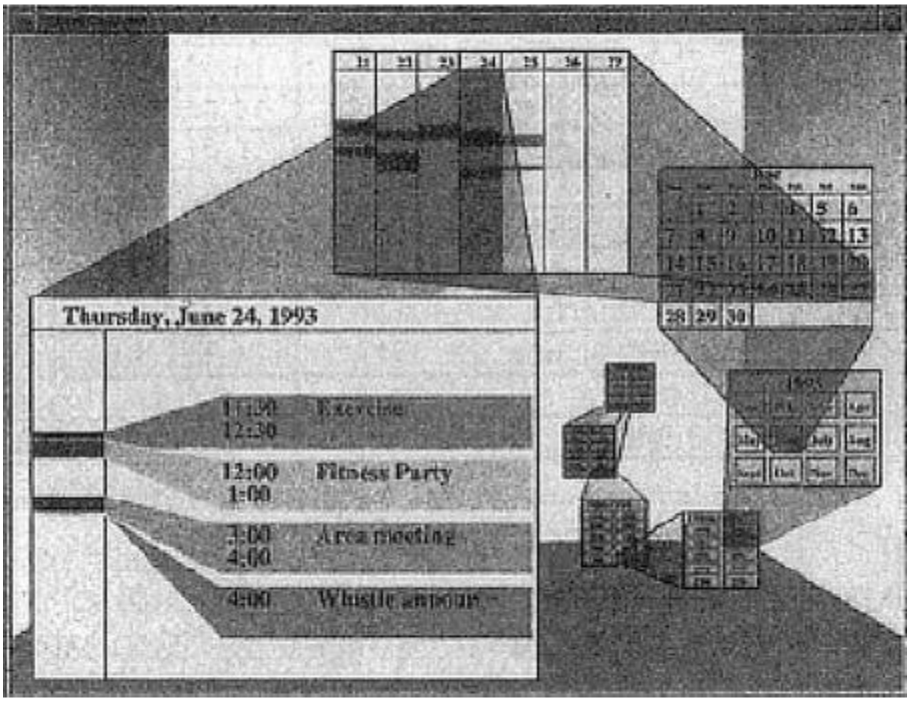


Figure 3.37: The Spiral Calendar (Zhai, Buxton, & Milgram, 1995)

The advantage of the Multiple Windows technique over the others listed previously is that this technique will not distort the object. It is able to show the focus and the context simultaneously.

Summary

The foundation of the art of Data Visualisation is to map data onto geometry and colour. The challenge is to figure out what colour and shapes work best, where to put them, and how to size them.

Visual Cues, Coordinate Systems, Scales and Context are the components of data visualisation. Visual Cue is the main component that a Data Visualisation viewer will see, while Coordinate System and Scale provide structure and a sense of space to the data visualisation. Last but not least, Context make Data Visualisation understandable and relevant.

For example, when we use length as the Visual Cue, the Cartesian coordinate system and the Categorical scale on the horizontal axis (with the Linear scale on the vertical axis in the context of understanding a company's sales amount by financial year), will derive a bar chart of sales amount by financial year.

In summary, we need to know how these components of Data Visualisation work. Once we have completed our design, we can try to ask other people to look at our design and observe whether they can understand what they see.

Exercises

ACTIVITY 1

Prepare your answers for the following questions as your instructor will ask you to present your answers in front of the class:

1. Describe the four components of Data Visualisation.
2. Compare the differences among the eight components of Visual Cues.
3. Explain the Trichromatic Theory.
4. Compare the differences among CIE, RGB and HSV colour systems.
5. Illustrate the applications of Visual Cues in visualising business related data.
6. Compare the differences among Cartesian, Polar, and Geographic Coordinate systems.
7. Illustrate the applications of Coordinate Systems in visualising business related data.
8. Compare the differences among Linear, Algorithmic, Categorical, Ordinal, Percent, and Time scales.
9. Illustrate the applications of Scales in visualising business related data.
10. Illustrate the applications of Context in visualising business related data.
11. Explain what Focus-context problem is.
12. Compare the Distortion, Rapid Zooming, Elision, and Multi-Windows techniques to solve Focus-context problem.

ACTIVITY 2

Students are to split into their groups to work in group on Assignment 2.

References

Aquarius.Net. (2002). *Conic projections.* Retrieved 21 June 2014 from http://www.mgaqua.net/AquaDoc/Projections/Projections_Conic.aspx.

Bederson, B., & Hollan, J. (1994). Pad++: A zooming graphical interface for exploring alternate interface physics. *Proceedings of UIST '94, ACM*, 17–36.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think.* San Francisco, CA, USA: Morgan Kaufmann Publishers.

Cockburn, A., Karlson, A., & Bederson, B.B. (2007). *A review of overview+detail, zooming, and focus+context interfaces*. Retrieved 28 June 2014 from https://www.cs.ubc.ca/~tmm/courses/533/readings/cockburn-ComputingSurveys09.pdf.

Few, S. (2015). *Signal: Understanding What Matters in a World of Noise.* Burlingame, CA, USA: Analytics Press.

Friendly, M. (2009). *Gallery of data visualisation*. Retrieved 21 June 2014 from http://euclid.psych.yorku.ca/SCS/Gallery/images/tlens01-s.gif.

Furnas, G. W. (1986). Generalized fisheye views. *Proceedings of CHI '86, ACM*, 17–26.

Furuti, C.A. (2013). *Polyconic projections*. Retrieved 21 June 2014 from http://www.progonos.com/furuti/MapProj/Normal/ProjCon/ProjPolycon/projPolycon.html.

Gaba, E. (2008). *File: world map sinusoidal proj-odeg centered.svg*. Retrieved 21 June 2014 from http://commons.wikimedia.org/wiki/File:World_map_sinusoidal_proj-0deg_centered.svg.

Glen. (2014). *CIE Chromaticity Diagram*. Retrieved 14 June 2014 from http://www.ntsc-tv.com/ntsc-index-05.htm.

Heer, J. (2004). *Position paper for the workshop on information visualisation software infrastructures at IEEE 2004 visualisation*. Retrieved 21 June 2014 from http://vw.indiana.edu/ivsi2004/jherr/.

Helmholtz, H. L. F. (1866). *Handbuch der Physiologischen Optik*, 1st edn. (3rd edn. transalted by J. P. C. Southall, *Opt. Soc. Am.*, 1924-5.)

Knippers, R. (2009). *Map projections*. Retrieved 21 June 2014 from http://kartoweb.itc.nl/geometrics/map%20projections/mappro.html.

Krygier, J.mB. (2014). *The power of maps and GIS*. Retrieved 21 June 2014 from http://go.owu.edu/~jbkrygie/krygier_html/geog_222/geog_222_lo/geog_222_lo01.html.

Lamping, J., Rao, R., & Pirolli, P. (1995). A focus + content technique based on hyperbolic geometry for viewing large hierarchies. *Proceedings CHI '95, ACM*, 401–408.

MacEvoy, B. (2005). *Do "primary" colors exist?*. Retrieved 14 June 2014 from http://www.handprint.com/HP/WCL/color6.html.

Mapof.net. (2014). *World map winkel tripel proj – the world maps*. Retrieved 21 June 2014 from http://mapsof.net/map/world-map-winkel-tripel-proj#.VAM6q_zlrIV.

Rao, R., & Card, S. K. (1994). The table lens: Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. *Proceedings of CHI '94, ACM*, 318–322.

Rogowitz, B., & Treinish, L. (1996). *Isomorphic color map*. Retrieved 21 June 2014 from http://www.research.ibm.com/people/l/lloydt/color/color.htm.

Sears, T. (2014). *Black bear population boom*. Retrieved 14 June 2014 from http://thejunglestore.blogspot.sg/2013/07/black-bear-population-boom.html.

Stockman, A. (2003). *Colorimetry*. Institute of Ophthalmology. London, UK.

The MathWorks Inc. (2014). *Converting color data between color spaces*. Retrieved 14 June 2014 from http://www.mathworks.com/help/images/converting-color-data-between-color-spaces.html.

The University of Mexico. (1996). *Zoomable graphical sketchpad*. Retrieved 21 June 2014 from http://www.cs.umd.edu/hcil/pad++/papers/jvlc-96-pad/zoom-5.html.

TUBS. (2011). *File: world location map (equirectangular 180).svg*. Retrieved 21 June 2014 from http://commons.wikimedia.org/wiki/File:World_location_map_(equirectangular_180).svg.

Stam, J. (2007). *Simulating diffraction*. Retrieved 14 June 2014 from http://http.developer.nvidia.com/GPUGems/gpugems_ch08.html.

Valentine, A. (2013). *Map projections 2008*. Retrieved 21 June 2014 from http://aspie.blog-it.me.uk/krb-3-main-types-0f-map-projections.phtml.

Waterman, S. (1999). *The world of maps*. Retrieved 21 June 2014 from http://www.watermanpolyhedron.com/worldmap.html.

Worthey, J. A. (2010). *Color matching with amplitude not left out*. Retrieved 14 June 2014 from http://www.jimworthey.com/jimtalk2004nov.html.

Yau, N. (2013). *Data points: visualization that means something.* Indianapolis, IN, USA: John Wiley & Sons Inc.

Young, T. (1802). On the theory of light and colours. *Phil. Trans. R. Soc. Lond.* 92, 20-71.

Zhai, S., Buxton, W., & Milgram, P. (1995). *The partial occlusion effect: utilizing semi-transparency in 3D human computer interaction*. Retrieved 28 June 2014 from http://etclab.mie.utoronto.ca/people/shumin_dir/SILK/silk.html.

STUDY UNIT 4
BASIC DATA VISUALISATION TECHNIQUES

Learning Outcomes

At the end of this unit, you are expected to:

- explain the best practices of visualise categorical data
- create bar, stacked bar, and side-by-side bar charts using Data Visualisation software
- create pie charts and area-fill charts using Data Visualisation software
- create heat maps and tree maps using Data Visualisation software
- explain the best practices to visualise time series data
- create line and spark-line charts using Data Visualisation software
- create Gantt chart using Data Visualisation software
- create trend and reference lines to charts using Data Visualisation software

Overview

This unit introduces several techniques to building Data Visualisation for categorical and time series data, and apply these techniques using Data Visualisation software.

1. Visualisation for Categorical Data

WATCH

Visualisation for Categorical Data
(Access video via iStudyGuide)

We have learnt in Study Units 2 and 3 that designing good data visualisation is surely an art, but just as surely, it is one that ought to be supplemented by science. Over the last decade, there has been modest revolution in the analysis of categorical data. Graphical methods and techniques of data visualisation are so commonly used for quantitative data that the development for such technique to be applied to frequency data and discrete data has begun.

1.1. Best Practices to Visualise Categorical Data

One-way frequency tables may be conveniently displayed in a variety of ways. Typically as bar charts which are often ordered by frequency, rather than by bar-label, dot charts (Cleveland, 1993), or pie charts when the percentage of the chart expression's total is important. For two- and higher-way tables, we can show the observed frequencies in the cells in relation to what we expect those frequencies to be under a reasonable null model. For example, the hypothesis that the variables of the row and the column are unassociated.

A Sales Manager at an apparel company like Charles and Keith in Singapore may have shoes sales data as shown in Figure 4.1 below. The figure shows the shoes sales data on the various categories of shoes for the past one year.

Shoe Category	Sales($'000)
Ballerinas	73
Boots	48
Flats	23
Heels	70
Loafers	35
Mary Janes	86
Mules	82
Pumps	36
Sneakers	18
Wedges	56

Figure 4.1: Shoe sales data

It will be challenging to imagine how the sales trend looks like, if the Sales Manager only use the above data presentation. Thus, the Sales Manager can use a bar chart to visualise the above data shown in Figure 4.2. Each rectangle represents a category. The longer the rectangle is, the greater the value it represents. Whether a higher value means better or worse depends on the dataset and the analysis of the point of view.

In this example, the higher the value is the better the sales performance is. We can see that the bar graphs are restricted in a way that each bar must start at the zero-axis, and must extend straight across or upward to the corresponding value. Do take note that we should avoid using a line chart to depict categorical data.

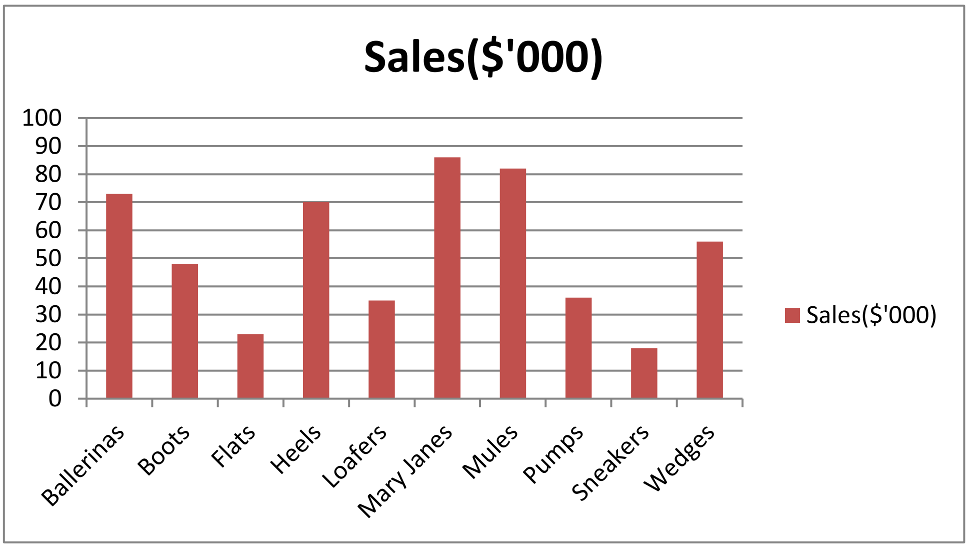


Figure 4.2: Using bar chart to visualise the shoe sales data

When we put categorical data together, the sum of the parts can equal to a whole. Using the same business scenario, a Sales Manager would like to view the sales data for each shoe category versus the aggregate sales amount. Visualising categories as a single unit can be beneficial if the Sales Manager wants to see distributions or the spread across a single population. This is where the pie chart, as illustrated in Figure 4.3, comes into the picture.

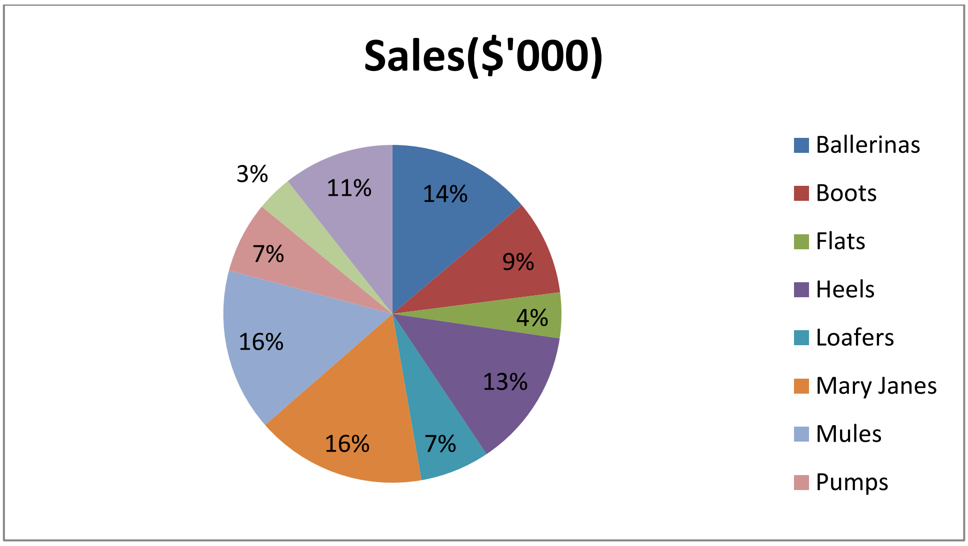


Figure 4.3: Using pie chart to visualise ten-year sales data

A full circle represents 100 percent, and each wedge is a portion of that 100 percent. The sum of all wedges equals 100 percent, and the angle is the visual cue. Pie charts tend to clutter quickly when we have a lot of categories, simply because there is only so much space in the circle and small values end up as slivers.

Sometimes, we may encounter categorical data within categories. This type of categorical data is called *sub-categorical data*. They are often more revealing than main categorical data. As we drill down from main category to sub category, there may be higher variability and more interesting insight to visualise.

Using the same business scenario, a Sales Manager would like to visualise the sales performance on per financial year basis, and he would like to further drill down to the product that has the highest sales amount in a particular financial year. He can use a Tree-map chart, as illustrated on Figure 4.5, to visualise the sub-categorical data within the categorical data.

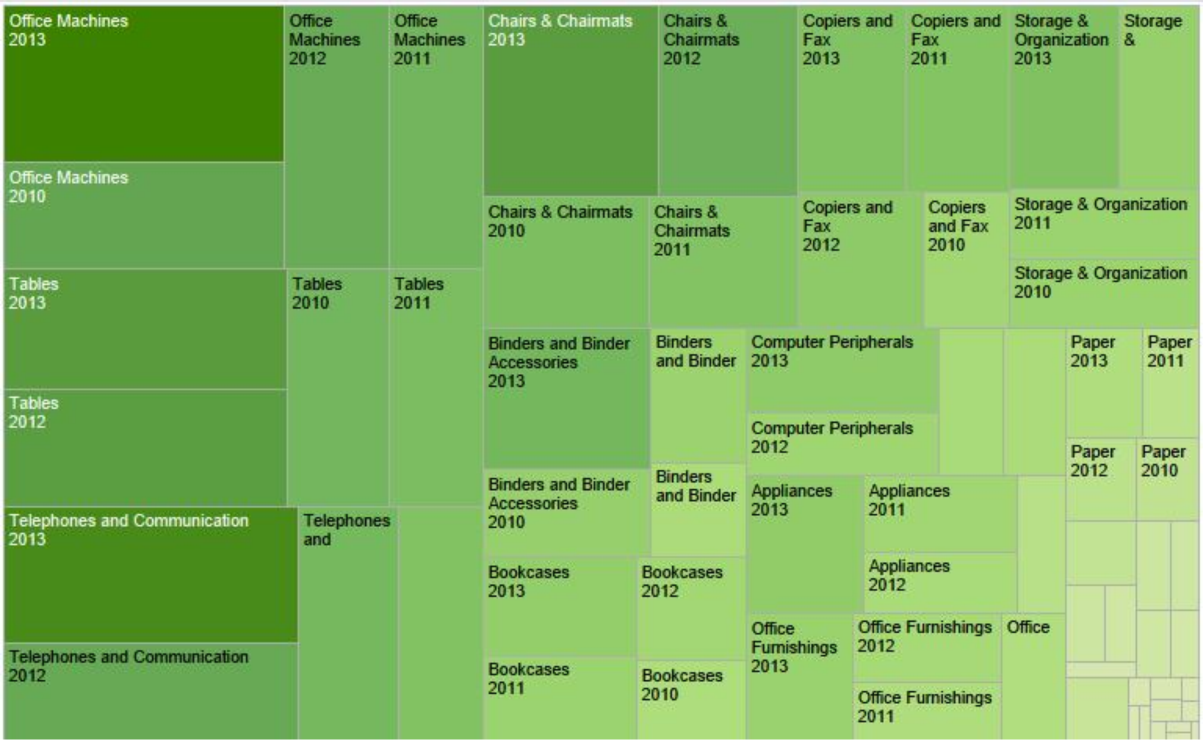


Figure 4.5: Using Tree-map chart to visualise sales data by year category and product sub-category

With categorical data, we often look for both maximum and minimum values because they give us a sense of the data range, and they can be easily found by sorting the values. After looking for both maximum and minimum values, we shall then look at the distribution of the data, i.e. whether it is positively skewed, negatively skewed, or normally distributed. Finally, we need to look for structure and patterns. If several categories have the same or different value, it is worth asking why and what makes the categories similar or different.

1.2. Creating Charts to Visualise Categorical Data

According to the Tableau Online Help, to create charts in the Tableau Desktop, we need to first open a workspace page. There are many ways we can open a workspace page. For example, we can drag and drop any data source icon (i.e. excel file icon) onto the Tableau icon. It will open the Tableau worksheet page for the selected data source. We can open as many connections as we want in Tableau by going to the start page or data connection page to select a new connection. Figure 4.6 shows the Tableau worksheet that is connected to a sample superstore sales excel data set used to create scatter plots.

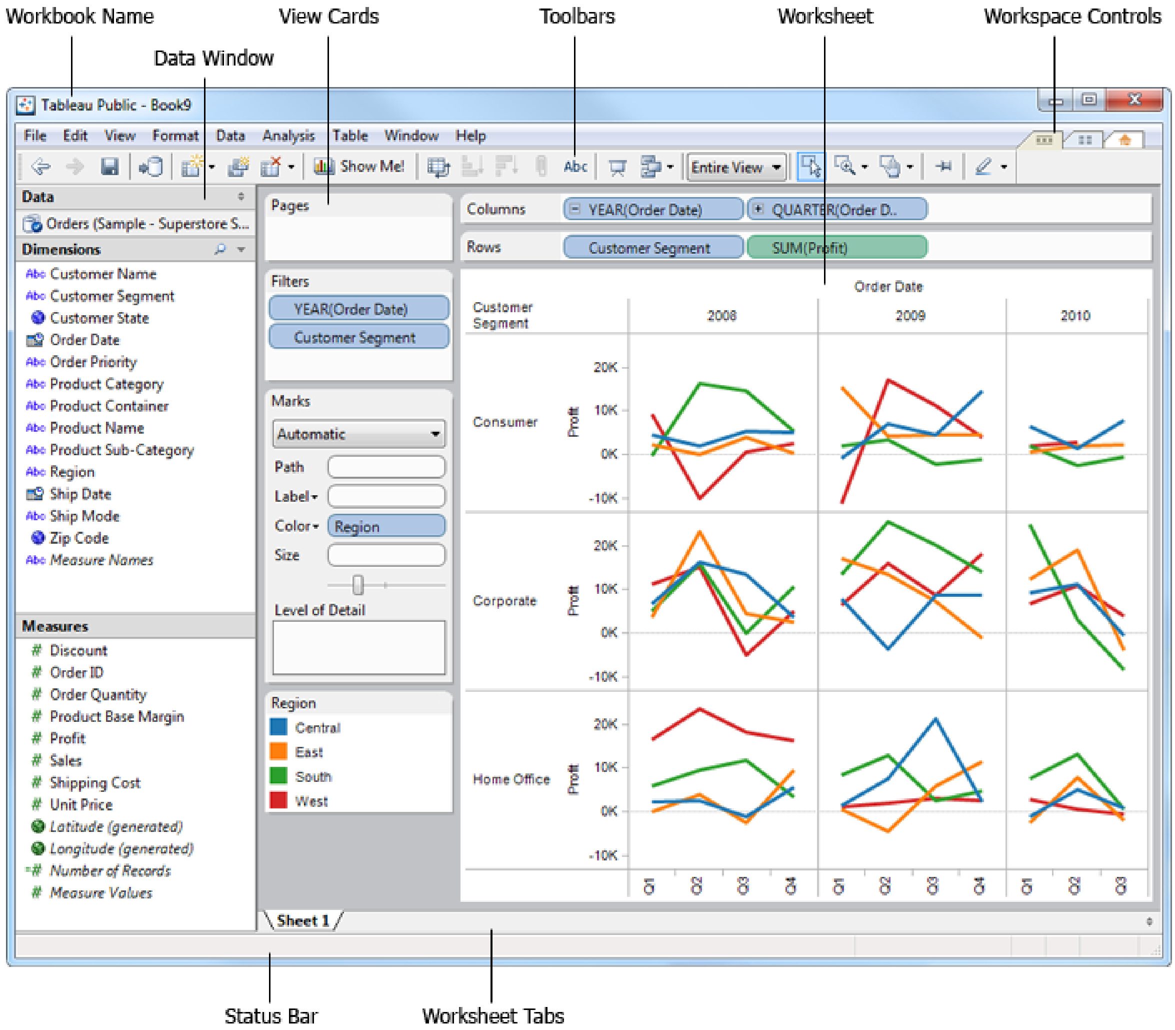


Figure 4.6: Tableau's Workspace (Tableau Online Help, 2014)

Data sources appear on the left side of the workspace in the Data window, as shown in Figure 4.7. We can hide and show the Data window by selecting **View -> Data Window**. Alternatively, we can also click the minimise button in the upper right corner of the Data window.

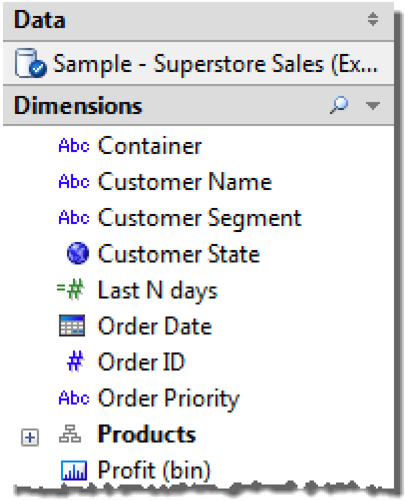


Figure 4.7: Tableau's Data Window (Tableau Online Help, 2014).

We can search the fields in the Data window by clicking the magnifying glass icon, and typing in the text box. To access commands related to each field, we can right click the field in the data window.

The Tableau's toolbar contains commands such as Connect to Data and Save, as shown in Figure 4.8. In addition, the toolbar contains analysis and navigation tools such as Sort, Group, and Zoom. We can unlock the toolbar by grabbing its left edge and then dragging it to a new location. We can hide or display the toolbar by selecting **View -> Toolbar**.

Figure 4.8: Tableau's Toolbar

Figure 4.8 shows the functions of each toolbar button:

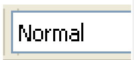
Toolbar Button	Description
	Undo: undoes the last task you completed.
	Redo: repeats the last task you cancelled with the Undo button.
	Save: saves the changes made to the workbook.
	Connect to Data: opens a dialogue box where you can create a new connection or select one from your repository.
	New Sheet: creates a new, blank worksheet.
	Duplicate Sheet: duplicates the existing content onto the new worksheet.
	Clear: clears the current worksheet. Use the drop-down list to clear specific parts of the view such as filters, formatting, and sizing.
	Automatic Updates: controls whether Tableau automatically updates the view when changes are made. Use the drop-down list to automatically update the entire sheet or just quick filters.
	Run Update: runs a manual query of the data to update the view with changes when automatic updates is turned off. Use the drop-down list to update the entire sheet or just quick filters.
	Show Me!: displays alternative views of the data, in addition to the best view according to best practices. The options available depend on the selected data fields when you click this button.
	Swap: moves the fields from the Rows shelf to the Columns shelf and vice versa. The Hide Empty Rows and Hide Empty Columns settings are always swapped using this button.
	Sort Ascending: applies a manual sort in ascending order of a selected field based on the measures in the view.
	Sort Descending: applies a manual sort in descending order of a selected field based on the measures in the view.
	Group Members: creates a group by combining selected values.
	Show Mark Labels: toggles between showing and hiding mark labels for the current sheet.
	Presentation Mode: toggles between showing and hiding everything but the view.
	View Cards: shows and hides the specified cards in a worksheet. Select the cards you want to hide or show from the drop-down list.
	Fit Selector: specifies how the view should be sized within the application window. Select either a Normal Fit, Fit Width, Fit Height, or Entire View.
	Fix Axes: toggles between locking the axes to a specific range and showing all of the data in the view.
	Highlight: turns on the highlighting function for the selected sheet. Use the options on the drop-down list to define how values will be highlighted.

Figure 4.9: Tableau's Toolbar Buttons

The Status Bar is located at the bottom of the Tableau Workbook. It displays descriptions of menu items, as well as information about the current view. For example, the status bar in Figure 4.10 shows that the view has 72 marks shown in one row and one column. It also shows that the Sum(Sales) for all the marks is $51.3M.

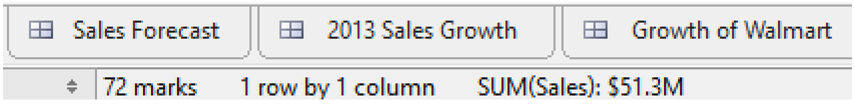


Figure 4.10: Tableau's Status Bar

We can hide the Status Bar by selecting **View -> Status Bar**.

Occasionally, the Tableau will display warning icons at the bottom right corner of the status bar to indicate errors that have or may occur. Figure 4.11 below shows the possible warning icons and what they mean.

Warning Icon	Description
1	Cancel Query Indicator: when you cancel multiple queries, an indicator will appear to show you how many queries are still running on the database and using the resources. For more information about this warning, refer to Abandoned Queries.
	Precision Warning: some fields are more precise in the database than Tableau can model. When you add these fields to the view, a precision warning will be displayed at the status bar. For more information about this warning, refer to Precision Warnings. Geocoding Warning: if Tableau cannot geocode some of your location values, this warning will show up. Geocoding warnings may happen if you have unknown location names or names that exist in multiple countries and states.

Figure 4.11: Tableau's Status Bar possible warning icons

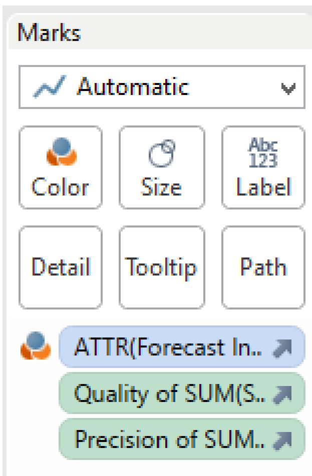


Figure 4.12: Tableau's Mark Card

Mark Cards can be shown or hidden, as well as rearranged around the worksheet. The following list describes each card and its contents.

- Columns Shelf is the shelf where we can drag fields to add columns to the view.
- Rows Shelf is the shelf where we can drag fields to add columns to the view.
- Pages Shelf is the shelf where we can create several different pages, with respect to the members in a dimension or the values in a measure.
- Filters Shelf is the shelf where we can specify the values to include in the view.
- Measure Names/Values Shelf is the shelf where we can use multiple measures along a single axis.
- Colour Legend contains the legend for colour encodings in the view. It is only available when there is a field on the colour shelf.
- Shape Legend contains the legend for shape encodings in the view. It is only available when there is a field on the shape shelf.
- Size Legend contains legend for the size encodings in the view. It is only available when there is a field on the size shelf.
- Map Legend contains the legend for symbols and patterns on a map. The map legend is not available to all map providers.
- Quick Filters is available for every field in the view. Use these cards to easily include and exclude values from the view without having to open the Filter dialogue box.
- Marks contains a mark selector where we can specify the mark type as well as the Path, Shape, Text, Colour, Size, Angle, and Level of Detail shelves. The availability of these shelves are dependent on the fields in the view.
- Title contains the title of the view. Double-click this card to edit the title.
- Caption contains a caption that describes the view. Double-click this card to edit the caption.
- Summary contains the summary of each of the measures in the view, including the Min, Max, Sum, and Average.

- Map Options allows us to modify the various labels and boundaries shown on the online maps. We can also use this card to overlay metro statistical area information.

- Current Page contains playback controls for the pages shelf, and indicates the current page that is displayed. This card is only available when there is a field on the pages shelf.

Each card has a menu that contains common controls that apply to the contents of the card. For example, we can use the card menu to show and hide the card. Access the card menu by clicking on the arrow at the upper right corner of the card.

1.2.1. Bar Chart, Stacked Bar Chart, and Side-by-Side Bar Chart

These charts facilitate one-to-many comparisons. The Bar Chart is the most effective way to compare values across dimensions due to their linear nature that makes precise comparisons easy. The Stacked Bar Chart should not be used when there are many different dimensions, because too many colours will be plotted in each bar. The Side-by-Side Bar Chart provides another way to compare measure across dimensions on a single axis.

The Country Manager of a stationary company, like Evergreen Stationary in Singapore, would like to visualise how much discount given for each product category sold. He can use Horizontal Bar Chart or Vertical Bar Chart in Tableau as illustrated in Figures 4.13 and 4.14. Below are the steps to creating a Bar Chart in Tableau:

1. Create a new worksheet.

2. For Horizontal Bar Chart, drag "0" or more dimensions into the worksheet's rows. For Vertical Bar Chart, drag "0" or more dimensions into the worksheet's columns.

3. For Horizontal Bar chart, drag one or more measures into the worksheet's columns. For Vertical Bar chart, drag one or more measures into the worksheet's rows.

4. Choose Horizontal Bars as the chart type.

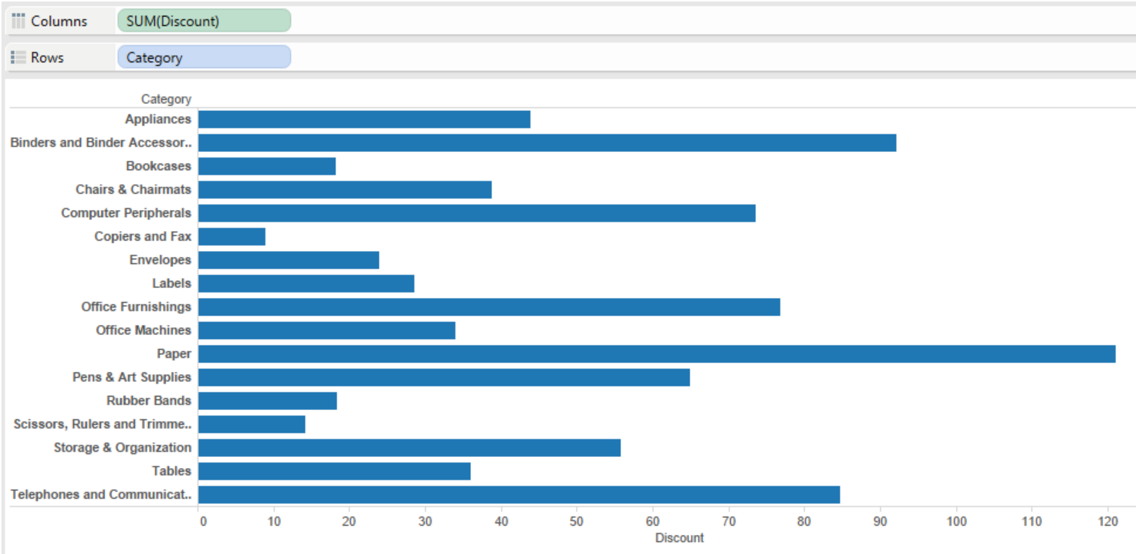


Figure 4.13: Horizontal Bar Chart setting

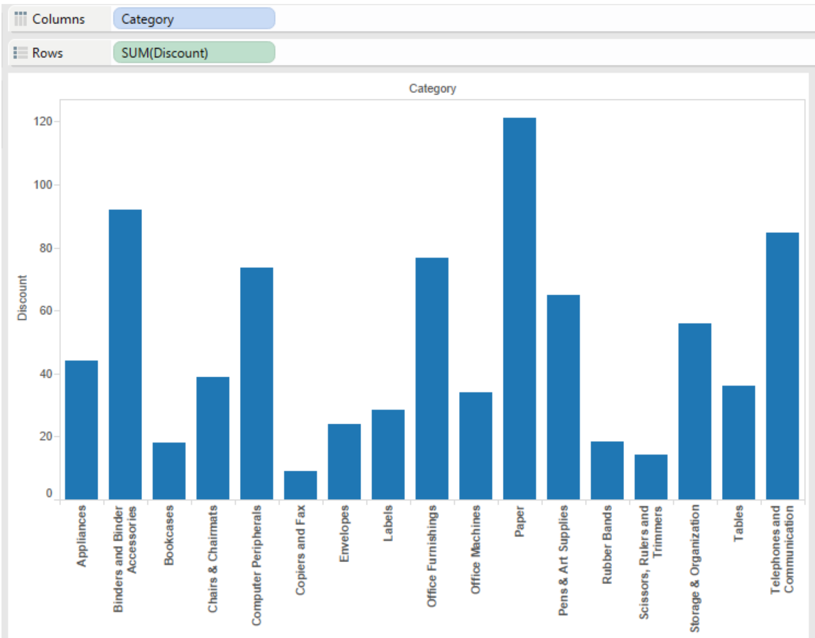


Figure 4.14: Vertical Bar Chart setting

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Murray. D. G. (2016), pp. 93-154.

2. Yau, N. (2013), pp. 97, 143.

3. Tableau Online Help - Bar Charts at http://onlinehelp.tableausoftware.com/v8.0/pro/online/en-us/buildexamples_bar.html.

The Procurement Manager of at a university like UniSIM in Singapore would like to visualise how much discount was given by suppliers for each product procured for the University. He can use Horizontal Stacked Bar Chart or Vertical Stacked Bar Chart in Tableau as illustrated in Figures 4.15 and 4.16. Below are the steps to creating a Stacked Bar Chart in Tableau:

1. Create a new worksheet.

2. For Horizontal Stacked Bar Chart, drag one or more dimensions into the worksheet's rows. For Vertical Stacked Bar Chart, drag one or more dimensions into the worksheet's columns.

3. For Horizontal Stacked Bar Chart, drag one or more dimensions, and one or more measures into the worksheet's columns. For Vertical Stacked Bar Chart, drag one or more dimensions, and one or more measures into worksheet's rows.

4. Choose Stacked Bars as the chart type.

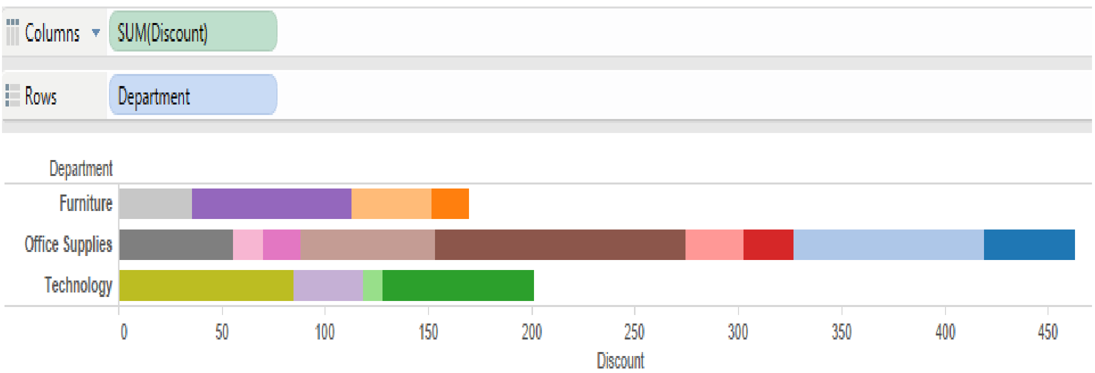


Figure 4.15: Horizontal Stacked Bar Chart setting

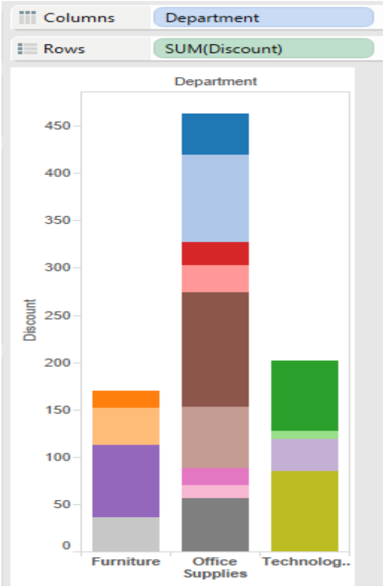


Figure 4.16: Vertical Stacked Bar Chart setting

READ

You should now read:

1. Yau, N. (2013), pp. 146.

2. Tableau Online Help – Stacked Bar Charts at http://kb.tableausoftware.com/articles/howto/stacked-bar-chart-multiple-measures.

The Sales Manager of an office supply and stationary company, like Evergreen Stationary in Singapore, would like to visualise the actual sales and forecasted sales for each product category. He can use Horizontal Side-by-Side Bar Chart or Vertical Side-by-Side Bar Chart in Tableau as illustrated in Figures 4.17 and 4.18. Below are the steps to creating a Side-by-Side Bar Chart in Tableau:

1. Create a new worksheet.

2. Drag two or more measures into either the worksheet's columns or rows. Measure Values and Measure Names will then be created automatically. Measure Names and Measure Values can be used to express various measures in a data set or multiple measures on a single axis. Measure Values contains the data while Measure Names is used to separate the bars used on the marks

card to distinguish the colour. Use the filters shelf to limit the number of measures shown in the view.

3. For Horizontal Side-by-Side Bar Chart, drag one or more dimensions and Measure Names into the worksheet's rows. For Vertical Side-by-Side Bar Chart, drag one or more dimensions and Measure Names into the worksheet's columns.

4. For Horizontal Side-by-Side Bar Chart, drag Measure Values into the worksheet's columns. For Vertical Side-by-Side Bar Chart, drag Measure Values into the worksheet's rows.

5. Choose Side-by-Side Bars as the chart type.

Figure 4.17: Horizontal Side-by-Side Bar chart setting

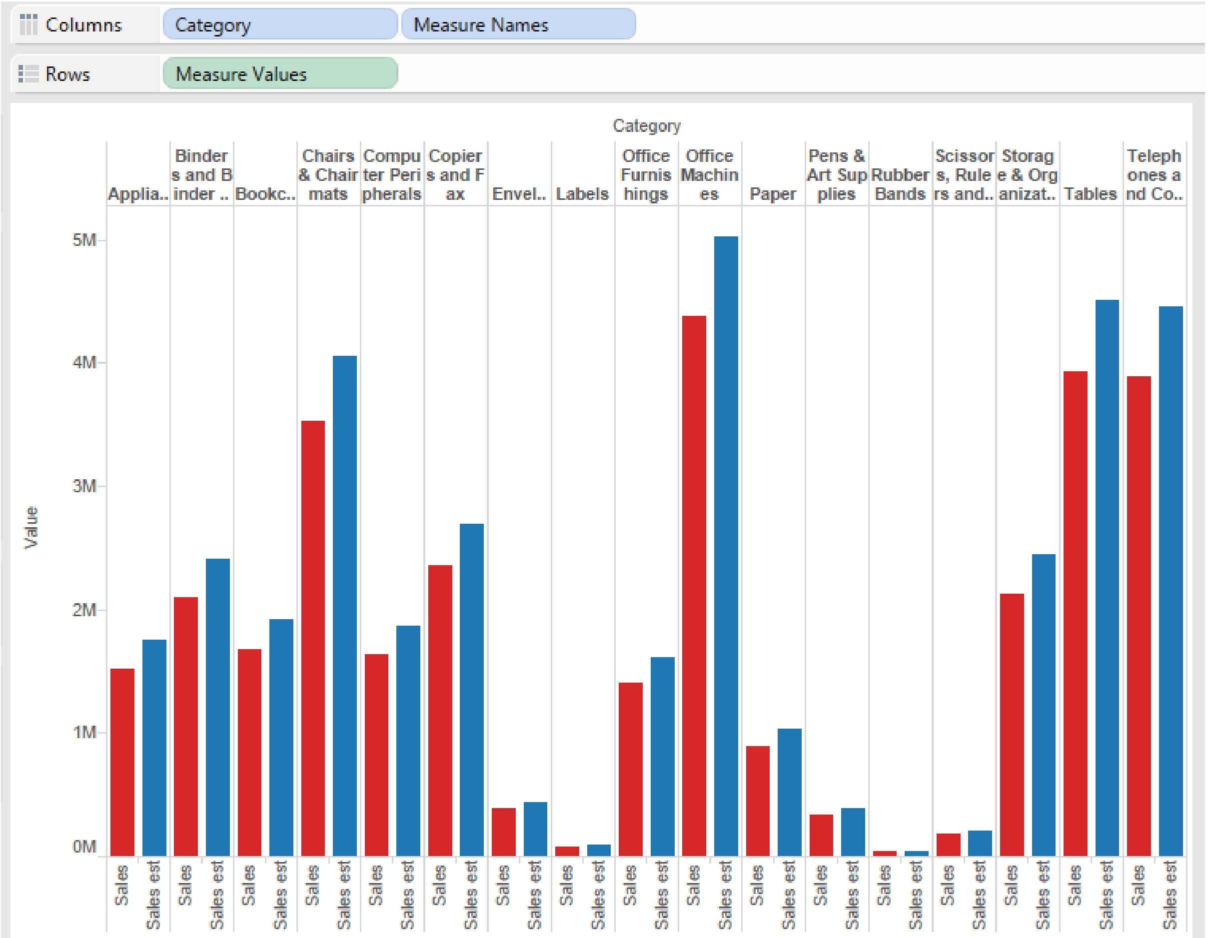


Figure 4.18: Vertical Side-by-Side Bar chart setting

READ

You should now read:

1. Tableau Online Help – Side-by-Side Bar Charts at http://kb.tableausoftware.com/articles/knowledgebase/grouped-bar-chart.

1.2.2. Pie Chart and Area Fill Chart

Area Fill chart plots values as bands, thus it is easy to misinterpret the top band as being the largest value in a data set. The Area Fill Chart is best used to plot a single dimension to avoid such misinterpretation. Pie Chart should be used to get a general sense of magnitude, but not for precise comparisons.

By limiting the Area Fill Chart to one dimension on each axis, and using a Pie Chart with only three slices, such combination is able to present information effectively. In addition, the Pie Chart can act as data filter, if we have limited space and we are sure that our Pie Chart's slices won't be too small.

Using the office supply and stationary company scenario, when the Sales Manager would like to visualise the distribution of each product category sales in year 2013. He can use Pie Chart in Tableau as illustrated in figures 4.19. Below are the steps to creating a Pie Chart in Tableau:

1. Create a new worksheet.

2. Drag one or more dimensions into the worksheet's rows, and one or two measures into the worksheet's columns.

3. Choose Pie Chart as the chart type.

4. Sort the dimensions in ascending or descending order by expressions.

5. Filter the expressions so that there is less number of slices.

6. Show "Mark Labels".

7. Edit "Chart Title".

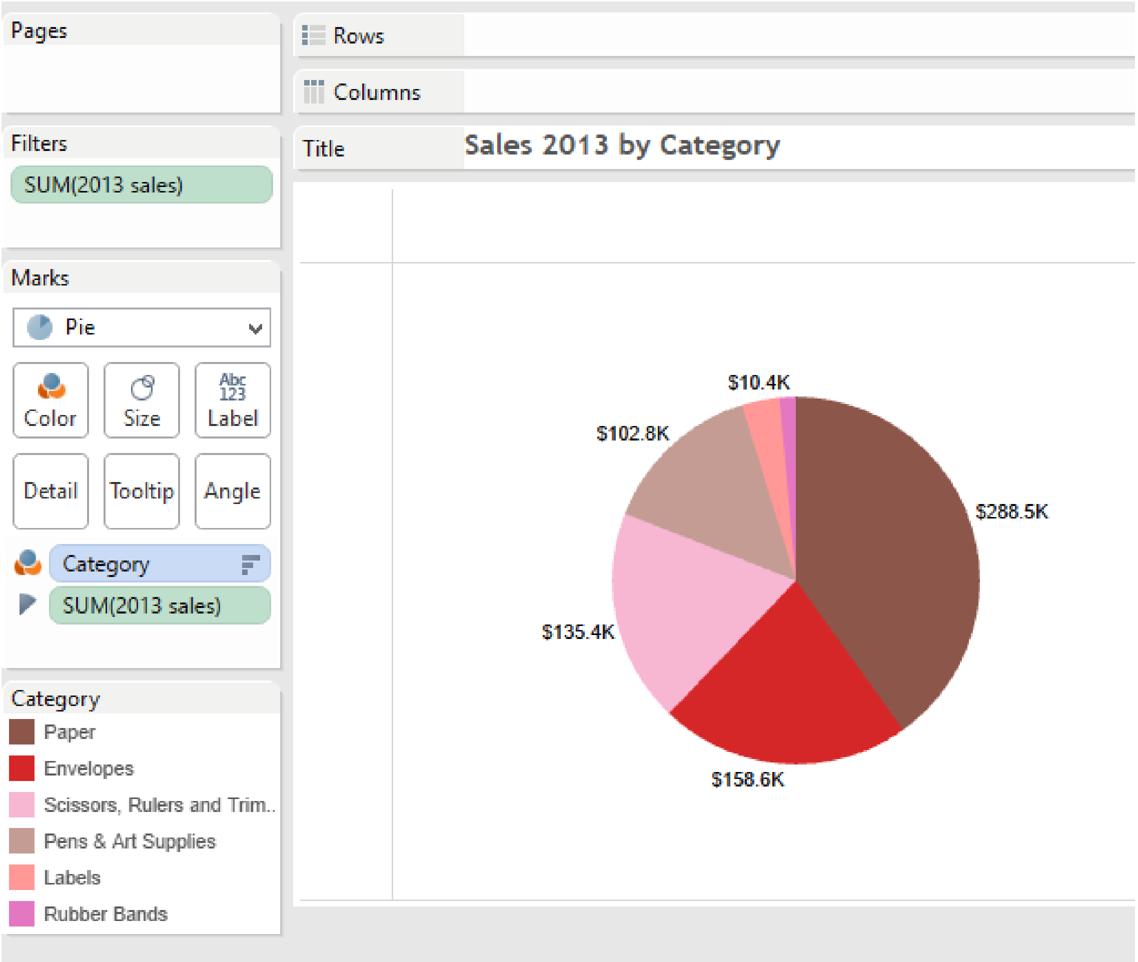


Figure 4.19: Pie Chart setting

You should now read:

1. Yau, N. (2013), pp. 98.

2. Tableau Online Help – Pie Charts at http://kb.tableausoftware.com/articles/knowledgebase/pie-chart.

Using the office supply and stationary company scenario, when the Sales Manager of the company would like to visualise the distribution of each product category sales in between year 2010 and 2013, instead of using Pie Chart, he can use Area Fill Chart in Tableau, as illustrated in figures 4.20. Below are the steps to creating an Area Fill Chart in Tableau:

1. Create a new worksheet.
2. Drag one date and "0" or more dimensions into the worksheet's columns.
3. Drag one or more measures into the worksheet's rows.
4. Choose Area Chart as chart type.
5. Sort the dimensions in ascending or descending order by expressions.
6. Show "Mark Labels".
7. Edit "Chart Title".

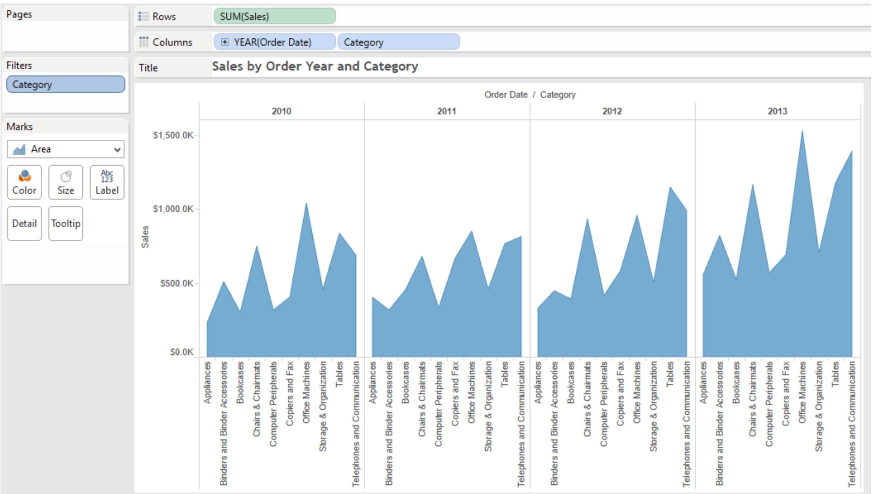


Figure 4.20: Area Fill Chart setting

READ

You should now read:

1. Tableau Online Help – Area Charts at http://downloads.tableausoftware.com/quickstart/feature-guides/area_charts.pdf.

1.2.3. Heat Map and Treemap

Comparing granular combinations of dimensions and measures can be done effectively with Heat Maps and Treemaps. Heat Maps use colours and sizes to compare up to two measures. Treemaps effectively display larger dimension sets using colours and sizes to display one or more dimensions, and up to two measures. In addition, these charts can also be used to replace Quick Filters on the dashboard.

Using the office supply and stationary company scenario, when the sales manager of the company would like to visualise the distribution of each product category profit using colours and sizes, he can use Heat Map in Tableau, as illustrated in Figure 4.21. Below are the steps to creating a Heat Map in Tableau:

1. Create a new worksheet.

2. Drag one or more dimensions into worksheet's rows.

3. Drag one or two measures into worksheet's columns.

4. Choose Heat Maps as the chart type.

5. Edit "Chart Title".

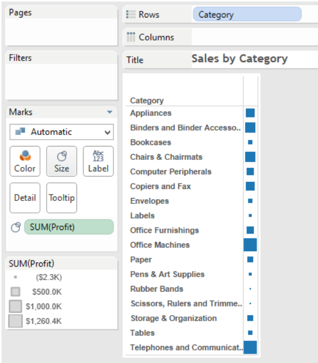


Figure 4.21: Heat map setting.

READ

You should now read:

1. Yau, N. (2013), pp. 164, 182-184, 195.

2. Tableau Online Help – Heat Maps Chart at http://onlinehelp.tableausoftware.com/current/pro/online/mac/en-us/help.htm#buildexamples_heatmap.html.

Using the office supply and stationary company scenario, when the Sales Manager of the company would like to visualise the distribution of each product category sales using colours and sizes, he can use Treemap in Tableau as illustrated in Figure 4.22. Below are the steps to creating a Treemap in Tableau:

1. Create a new worksheet.

2. Drag one or more dimensions into the worksheet's rows.

3. Drag one or two measures into the worksheet's columns.

4. Choose Treemaps as the chart type.

5. Edit "Chart Title".

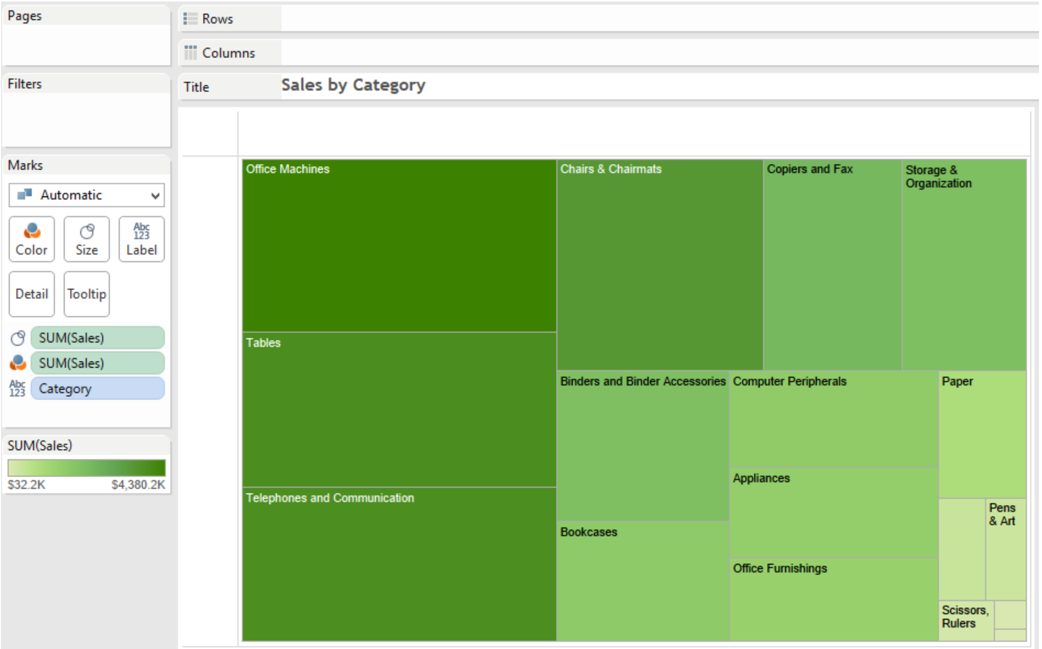


Figure 4.22: Treemap setting

You should now read:

1. Yau, N. (2013), pp. 53, 57, 146, 150, 281.

2. Tableau Online Help – Treemaps Chart at http://onlinehelp.tableausoftware.com/current/pro/online/mac/en-us/help.htm#buildexamples_treemap.html.

2. Visualisation for Time Series Data

WATCH

Visualisation for Time Series Data
(Access video via iStudyGuide)

We look at time every day. The most common thing we look for in time series or temporal data is trends, e.g. is something increasing or decreasing, is this a seasonal cycle. To find the patterns, we have to look beyond individual data points to get the whole picture. It is easy to pick out a single value from a point in time, but when we look at what came before and comes after, we can have a better understanding of what that single value means. Furthermore, the more we know about our data, the better the insight that we can gain.

2.1. Best Practices to Visualise Time Series Data

Time Series Data can be categorised as either discrete (bucketed) or continuous (unbroken). One consideration when presenting time series is how we treat the discrete and continuous time series data.

2.1.1. Discrete Time Series Data

For Discrete Time Series Data, values are from specific points or blocks of time, and there is a finite number of possible values. For example, the number of product sold last year is discrete. Last year's sales figure will not change afterward. Something like currency exchange rate is continuous, because it can be measured at any time of the day during any interval, and the rate is constantly changing.

A Line Chart is the most effective way to display time series data. A Line Chart for discrete time series data places breaks between time units like year, quarter, month and day.

Bar and Stacked Bar Charts can also be used to display time series data. The x-axis of the Bar and Stacked Bar Chart (or the time axis) provide a place for points in time that are ordered chronologically. The y-axis of the Bar and Stacked Bar Chart (or the value axis) indicates the scale of the graph. Always start the value axis at zero, otherwise the Bar or Stacked Bar Chart could display incorrect relationship.

Sometimes, points can also be used to display time series data instead of bars because points use less space and there are no bins. Points can also allow dashboard viewers feel that there is flow from one point to another. This type of point chart is commonly known as a Scatter Plot Chart. It can be used to display non-time series data. Scatter Plot Chart uses position as visual cue, so that we can compare each point in time to other points in time based on where they are placed in the X- and Y- coordinates. Because of this, the value axis of a Scatter Plot Chart does not always have to start at zero, even though it is usually a good practice.

2.1.2. Continuous Time Series Data

Visualising continuous time series data is similar to visualising discrete time series data. We still have a discrete number of data points, even if the dataset is continuous. The difference between visualising continuous time series data from visualising discrete time series data is in what they represent in the physical world. Continuous time series data represents a constantly changing phenomena.

Line Chart for continuous time series data is presented as unbroken lines. We have the points on X- and Y- coordinates, and the line will connect the points to help us see the trend in our data. It is usually a best practice to start the value axis at zero so that it will not affect the scale. How far we can stretch the X-axis can also affect the appearance of the trend. If we stretch too far out, we may not see the pattern. If we squish too much, the increase from point to point may look more than it is.

One of the drawbacks of a standard Line Chart is that it implies steady change from one point to other points. That is about right with a measure like country's population, but that will not be right for a case where a measure stays at a value for a long time and all of a sudden it declines or inclines. For example, interbank interest rate may stay the same for months and drop within a day. For such continuous time series data type, we need to use a Step Chart. Instead of directly connecting one point to another, the line stays at the same value until there is a change, at which point it jumps up or down to the next value.

When we have a lot of data, or the data we have is *noisy*, it can be hard to spot trends and patterns. Thus, it will be easier if we can estimate the trend line. To estimate the trend line, we can draw a line that goes through where there is the most points, and minimise the summarised distance from these points to the fitted line. The most straight forward method is to create a straight fitted line using the basic slope intercept equation below:

$$Y = mx + b$$

The "m" refers to the slope of the fitted line, "b" refers to the intercept of the fitted line to the Y-axis. When the fitted line is not linear, we can use a statistical method created by William Cleveland and Susan Devlin called LOESS, or Locally Weighted Scatterplot Smoothing. This method will enable us to fit a curve to our data, because LOESS starts at the beginning of the data and takes small slices. At each slice, it estimates a low-degree polynomial for just the data in the slice. LOESS moves along the data, fits a bunch of tiny curves, and together the tiny curves form a single curve.

You should now read:

1. Yau, N. (2013), pp. 154-161, 188.

2.2. Creating Charts to Visualise Time Series Data

To create charts in the Tableau, we first need to open a workspace page, and select the data connection. We shall use the sample of Superstore Sales data to illustrate the steps of building charts in this section.

2.2.1. Line Chart

Line Chart is the most effective way to display time series data. Discrete time series data are presented with breaks between time units, while continuous time series data are presented in unbroken lines.

The Sales Vice President of a food and beverages company like Burger King would like to visualise the sales trend in between year 2010 and 2013. He can use Line Chart in Tableau as illustrated in Figure 4.23. Below are the steps to creating a Line Chart in Tableau:

1. Create a new worksheet.

2. Drag one time series data and "0" or more dimensions into the worksheet's columns.

3. Drag one or more measures into the worksheet's columns.

4. Choose Lines (Continuous) for continuous time series data, or Lines (Discrete) for discrete time series data as the chart type.

5. Edit "Chart Title".

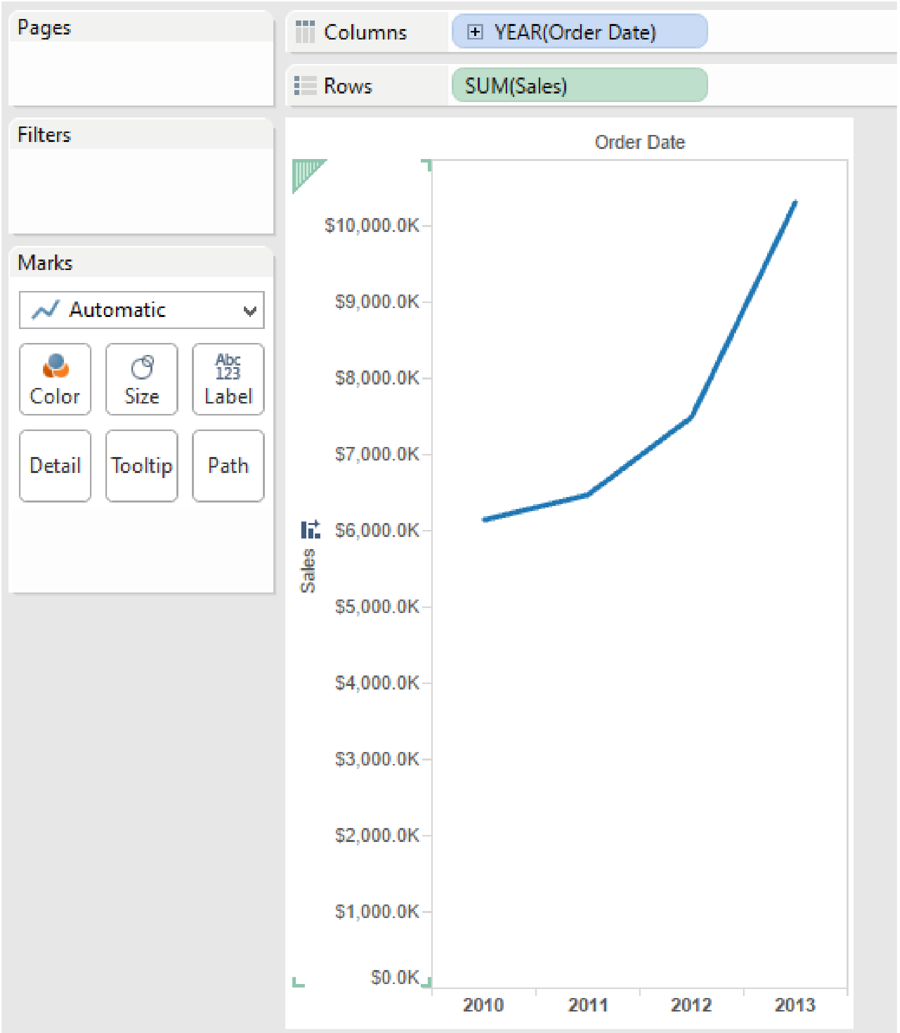


Figure 4.23: Line Chart setting

READ

You should now read:

1. Yau, N. (2013), pp. 155, 159.

2. Tableau Online Help – Line Charts at http://onlinehelp.tableausoftware.com/current/pro/online/mac/en-us/help.htm#buildexamples_line.html.

2.2.2. Sparkline Chart

Sparkline Chart provides very effective time series data visualisation, especially when we have space constraint. Sparkline Chart will take up less space than the Tableau's default time series charts, like a Line Chart.

Using the food and beverages company scenario, when the Sales Vice President of the company would like to visualise the sales trend in between years 2010 and 2013 for various countries, he can use a Sparkline Chart in Tableau, as illustrated in Figures 4.24 and 4.25. Below are the steps to creating a Sparkline Chart in Tableau:

1. Create a new worksheet.

2. Drag one time series data and "0" or more dimensions into the worksheet's rows.

3. Drag one or more measures into the worksheet's columns.

4. Choose Lines (Continuous) for continuous time series data, or Lines (Discrete) for discrete time series data as the chart type.

5. Left click on each axis to select the axis that we would like to edit.

6. Right click on the axis and select "Edit Axis".

7. Make each axis range independent.

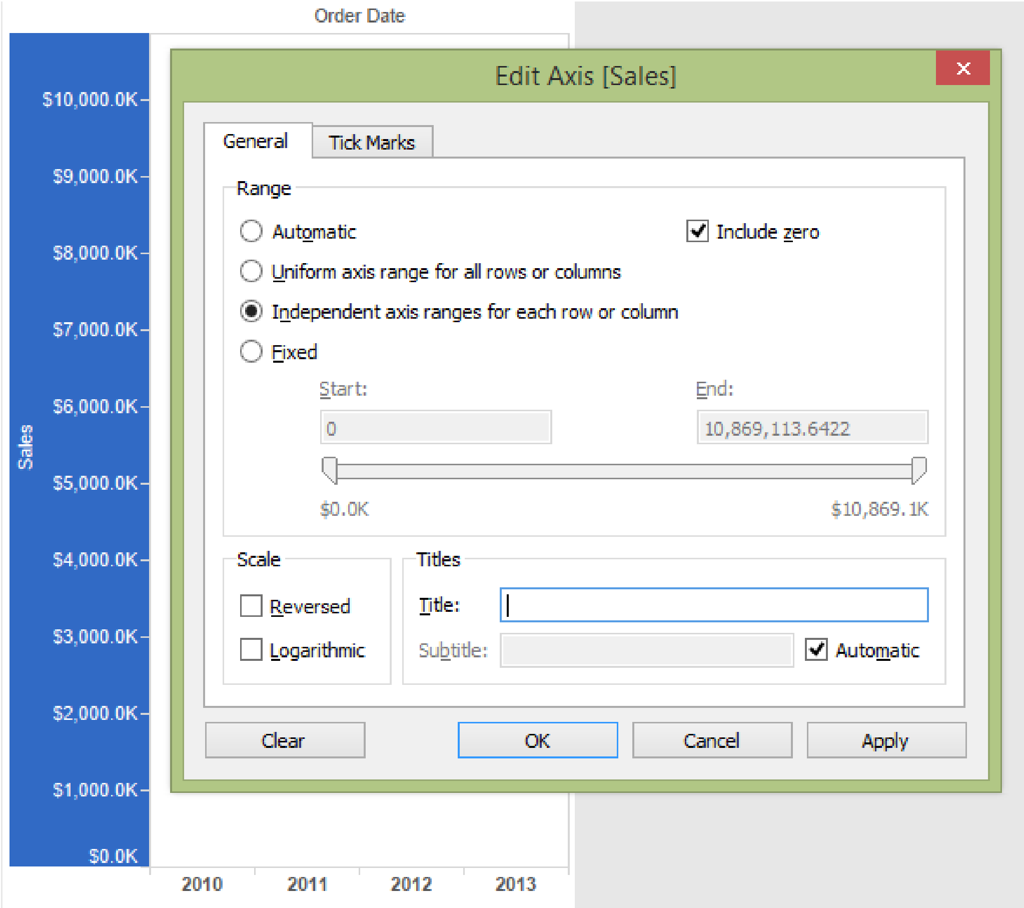


Figure 4.24: Sparkline Chart setting – independent axis range

8. Remove the axis title.

9. Drag the right edge of the chart to the left.

10. Drag the chart from bottom up.

11. Reduce the mark size from the Marks Card.

12. If necessary, emphasise the change using a table calculation.

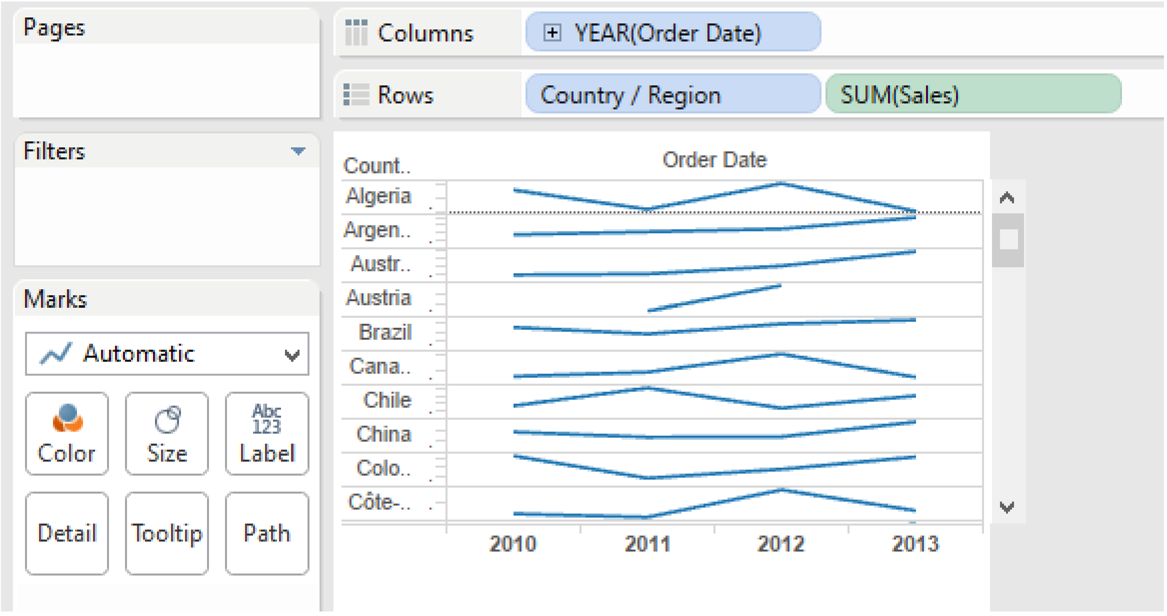


Figure 4.25: Sparkline Chart

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Murray. D. G. (2016), pp. 280.

2. Tableau Online Help – Sparkline Charts at http://kb.tableausoftware.com/articles/knowledgebase/sparklines-bullets.

2.2.3. Gantt Chart

A Gantt Chart is commonly being used for project planning. Activity or event start time is visualised by the bar's horizontal position, and the duration of each activity or event is visualised by the individual bar length. Thus, this chart is useful to visualise the timing and duration of activities or events.

The Sales Operations Manager of an e-commerce company like Amazon.com would like to visualise the order lead time for each order. He can use Gantt Chart in Tableau as illustrated in figures 4.26. Below are the steps to creating a Gantt Chart in Tableau:

1. Create a new worksheet.

2. Select "Analysis", then "Create Calculated Field".

3. In the “Calculated Field” dialogue box, type “Duration in Days” in the name textbox and type “DATEDIFF” ('second',[Order Date],[Ship Date])/86400.

4. Drag the “Order Date” field into the worksheet’s columns.

5. Right click on the “Order Date” field and select “Exact Date”.

6. Drag the “Order ID” field into the worksheet’s rows and into the “Colour Mark”.

7. On the Mark Card’s drop down menu, select “Gantt Bar”.

8. Drag the “Duration in Days” field to “Size”.

9. If there is any null value in “Order ID”, right click “Order ID” field and select “Hide” to hide the null value.

10. Right click on the “Order Date” field and select “Edit Axis”.

11. In the “Edit Axis” dialogue box, under “Range” select “Fixed”. Change the date range, and click “Apply”.

12. On the “Tick Marks” tab, fix the major and minor tick marks every one day. Change the origin date to 1 June 2010. When finished, click “OK”.

13. Drag “Order Date” field into Filters Cards to filter the “Order Date” from 1 June 2010 to 1 July 2010.

14. If necessary, we can edit the colours of the bars. Right click on the “Order ID” Legend Card and select “Edit Colour”. Select the preferred colour from the “Select Colour Palette” drop down menu. Under “Select Data Item”, select the item that we want to assign a different colour, and click on the colour on the right pane. Repeat this step for each item we want to assign a different colour. When finished, click “Apply” and “OK”.

Figure 4.26: Gantt Chart

You should now read:

1. Tableau Online Help – Gantt Charts at http://kb.tableausoftware.com/articles/knowledgebase/gantt-chart.

2.2.4. Trend and Reference Lines

When we try to visualise granular data, sometimes it may result in random looking data visualisation. Trend Lines can help us see patterns that can provide predictive value, by drawing a line that best fits the values in the visualisation. Reference Lines can help us compare the actual data presented in the visualisation against targets, create statistical analyses of the deviation contained in the visualisation, or create the range of values based on fixed or calculated numbers. By looking at Reference Lines, we will be able to identify outliers that may require our attention or additional analysis.

Using the food and beverages company scenario, when the Sales Vice President of the company would like to see the pattern of the sales in between years 2010 and 2012, he can use Trend Lines together with Line Chart in Tableau as illustrated in Figures 4.27. Below are the steps to creating Trend Lines in Tableau:

1. Create a chart, for example a line chart, by following steps that has been described in the Line Chart section previously.

2. Right click on the chart and select "Show Trend Lines".

3. For more the "Trend Lines" option, point at the "Trend Line", right click, and select "Edit Trend Lines". This will expose the Trend Line Menu. Select "Show Confidence Bands".

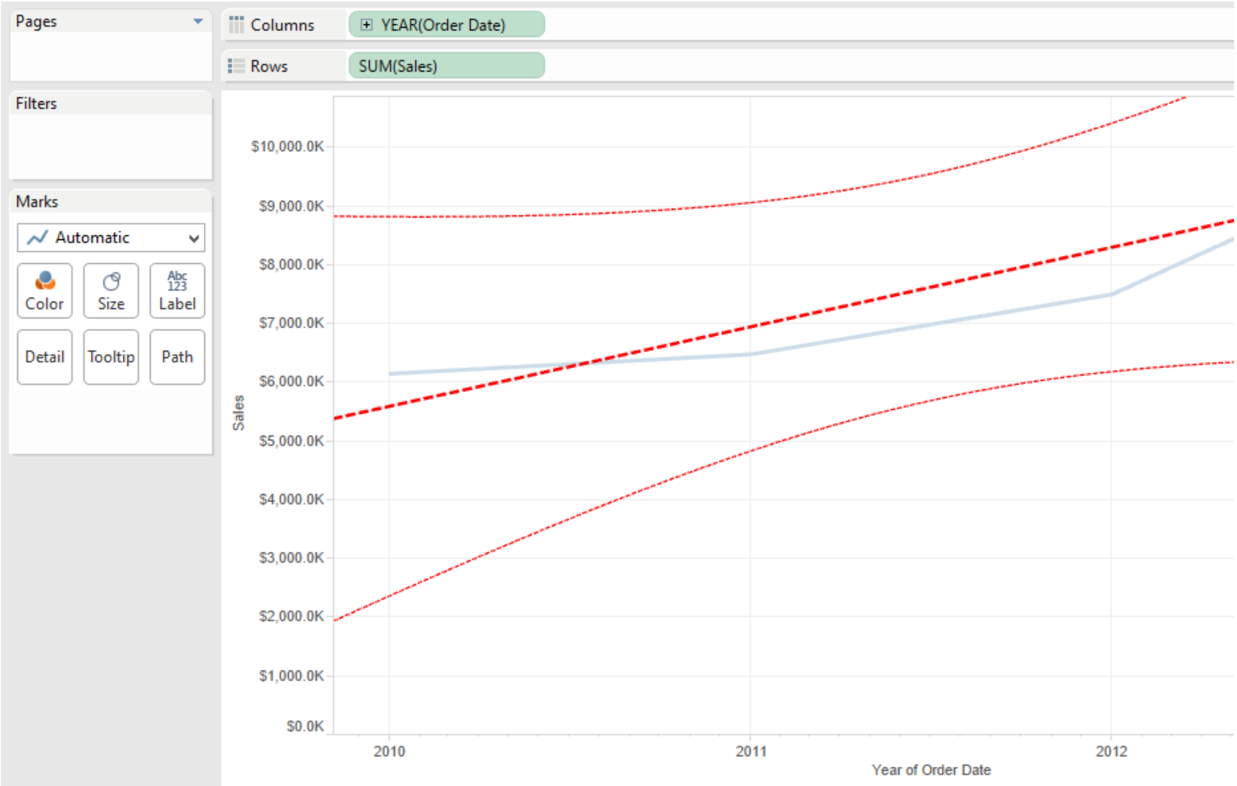


Figure 4.27: Trend Line

Using the food and beverages company scenario, when the Sales Vice President of the company would like to compare the actual sales against average sales in between years 2010 and 2012, he can use Reference Lines together with Line Chart in Tableau as illustrated in Figure 4.28. Below are the steps to creating Reference Lines in Tableau:

1. Create a chart, for example, a Line Chart, by following steps that has been described in the Line Chart section previously.

2. Right click on the axis at which we want to apply the reference line, and select "Add Reference Line", "Band", or "Box".

3. Explore the line, band, and distribution buttons in conjunction with the computation value's dropdown menu to see all the available options for reference line types.

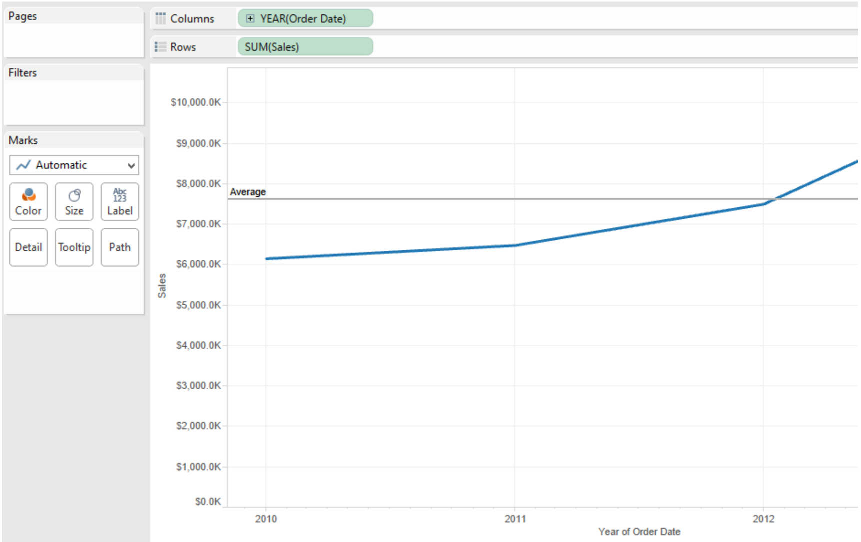


Figure 4.28: Reference Line

You should now read:

1. Tableau Online Help – Reference Lines at http://kb.tableausoftware.com/articles/knowledgebase/Reference-Lines.

2. Tableau Online Help – Trend Lines at http://onlinehelp.tableausoftware.com/v8.0/pro/online/en-us/trendlines_add.html.

Summary

In our daily life, more often than not, we may like to group our business related data based on customer, region and other dimensions, and store these data as categorical data. One of the most common charts to visualise categorical data is the Bar Chart, including Side-by-Side Bar Chart and Stacked Bar Chart.

We can use a Pie Chart to visualise data distributions across a single population when we only have one measure, or use a Stacked Bar Chart when we have more than one measures. Sometimes, we may encounter categorical data within categories. We can use a Tree-map Chart to visualise the sub-categorical data within the categorical data.

Time is embedded in our day-to-day life, thus visualising time series data can be quite intuitive. By visualising time series data, we can understand how thing changes and evolves, and the challenge is to find out how much the change or evolution is, and to know what to look for in the visualisation.

In general, it is easy to spot the trend on a time series data visualisation, and to say that something is good or bad, as that is what data visualisation is meant for. Nevertheless, we should take this further by drilling down on sections of time and questioning ourselves why were there drops or spikes on some days, but nowhere else.

After we understand our data well, we need to explain and highlight the insight so that data visualisation viewers will know which part of the visualisation they should pay more attention to. Data visualisation can be nice for its creator, but without context the visualisation may be meaningless for everyone else.

Exercises

ACTIVITY 1

Prepare your answers on the questions below as your instructor will ask you to present you answers in front of the class:

1. Briefly explain the best practices to visualise categorical data with examples.

2. Briefly explain the best practices to visualise time series data with examples.

3. Explain the use of trend and reference lines.

ACTIVITY 2

1. Using the country data and the Tableau workspace prepared in Week 2, create:

   a. Bar Chart, Side-by-Side Bar Chart, and Stacked Bar Chart

   b. Pie Chart and Area Chart

   c. Heat-map and Tree-map

   d. Line Chart, Sparkline Chart, Gantt Chart, Reference Line, and Trend Line

2. Instructor will randomly appoint students to present their work in front of the class.

Tableau Demos

Connect to Data

This demo aims to present how to import data into Tableau using Tableau Desktop.

At the end of this demo, you will be able to:
- use Connect To Data screen
- import data from excel, csv and text files
- connect to single and multiple tables
- use Live versus Extract (in-memory) feature
- create Charts to show the imported data

Click here to view the demo.

Heat Map and Treemap Charts

This demo aims to present how to create Heat Map and Treemap Charts using Tableau Desktop.

At the end of this demo, you will be able to:
- create Treemap Chart
- create Heat Map Chart

Click here to view the demo.

References

Cleveland, W. S. (1993). *Visualizing Data*. Summit, NJ, USA: Hobart Press.

Few, S. (2015). *Signal: Understanding What Matters in a World of Noise.* Burlingame, CA, USA: Analytics Press.

Murray, G. D. (2016). *Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software. (2nd Edition)* Indianapolis, IN, USA: John Wiley & Sons Inc.

Tableau Online Help. (2014). Retrieved 1 July 2014 from

http://onlinehelp.tableausoftware.com/current/pro/online/mac/en-us/help.htm#default.html,

http://onlinehelp.tableausoftware.com/v6.1/public/online/en-us/id10ANH0QH0XW.html

and

http://onlinehelp.tableausoftware.com/v6.1/public/online/en-us/id10ANG90H0RO.html.

Yau, N. (2013). *Data Points: Visualisation that Means Something*. Indianapolis, IN, USA: John Wiley & Sons.

STUDY UNIT 5
ADVANCED DATA VISUALISATION TECHNIQUES

Learning Outcomes

At the end of this unit, you are expected to:

- explain the best practice to visualise spatial data
- show spatial data on a map using Data Visualisation software
- prepare data to enable point-to-point mapping using Data Visualisation software
- explain the best practice to visualise multi-variable and distribution of data
- create scatter, circle and side-by-side circle plots using Data Visualisation software
- create bullet chart, bubble chart and histogram using Data Visualisation software
- create box plot and pare to chart using Data Visualisation software
- create parameters using Data Visualisation software
- create forecast using Data Visualisation software

Overview

This unit introduces several techniques in building Data Visualisation for ad-hoc analysis environment, and to apply the techniques using Data Visualisation software.

1. Visualisation for Spatial Data

WATCH

Visualisation for Spatial Data
(Access video via iStudyGuide)

Spatial data are related to the location of the subject matter, such as customer country, supplier address, etc. According to Yau (2013), there is a natural hierarchy to spatial data that allows, and often requires, us to explore at different granularities. For example, from the corporate level, the senior management may be interested to analyse the data based on country level. However, in order to know what happened in the business units, tactical and operational managers may want to zoom in on the data into provinces or states, counties, districts, cities, towns, neighbourhoods, all the way down to the individual household.

1.1. Best Practices to Visualise Spatial Data

The most common way to visualise spatial data is with maps that place values within a geographic coordinate. However, a map is not always the most informative way to visualise spatial data. For example, we can treat "Region" as a category and use a Bar Chart to visualise the data.

We can visualise the geographic coordinate of a location by mapping the latitude and longitude coordinates to two-dimensional space, and draw a point on the space. This seems quite straightforward, but it can pose challenges when there are a lot of locations to be visualised.

When the density of individual locations across a region is more informative than the overlapping points on a map, we may want to colour code the region based on the density scale, or use lines to show data continuously over geography. We can also size the regions by the data and ignore the physical area, so that the regions with high density data will appear bigger than the regions with low density data.

Rather than separating locations, we may sometimes want to explore connections between two locations. For example, when we analyse social media data, we would like to know the relationship between one person and the others. In this scenario, we can plot each individual on a map, and draw lines to connect one individual with the others.

1.1.1. Map

To create map in the Tableau, we need to assign "Geographic Role" to the field that contain geographic data, such as country, region, state, zip code, and so on. Fields with a Geographic Role will automatically generate Longitude and Latitude coordinates to be visualised as a map. Geographic roles are arranged in a hierarchy where the field with the lowest level of detail is what is shown on the map.

The Customer Relationship Manager of an e-commerce company like Amazon.com would like to visualise where their customers stay. He can use Map in Tableau as illustrated in Figure 5.1. Below are the steps to creating a Map in Tableau:

1. Create a new worksheet.

2. If necessary, assign the Geographic Role to fields that are not automatically recognised by Tableau as Geographic data.

3. Double click the "Region" field under Dimension of the Data window. The Tableau will automatically add Longitude and Latitude coordinates to the columns and rows shelves. The "Region" field is automatically placed on the Level of Detail shelf.

4. Drag the "Postal Code" field to the Level of Detail shelf to specify more details.

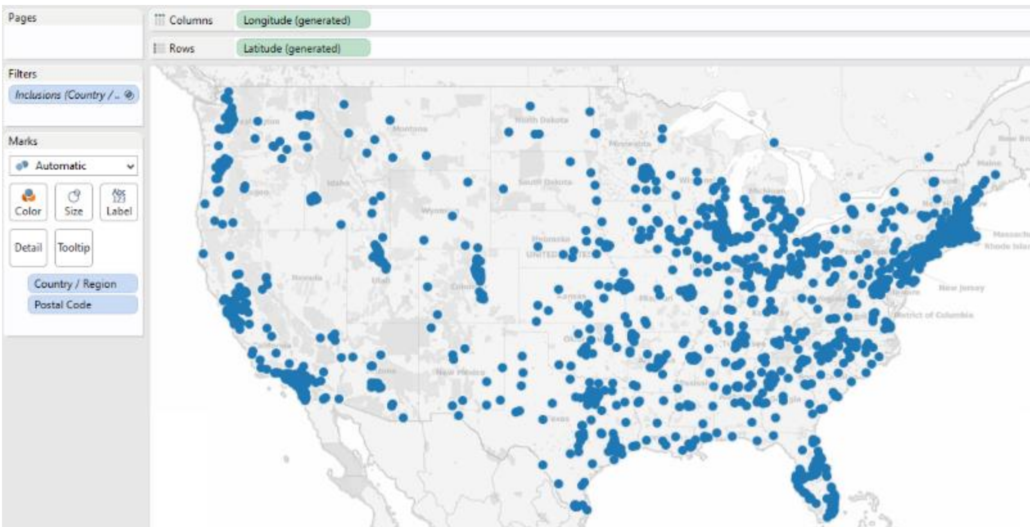


Figure 5.1: Map

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Murray. D. G. (2016), pp. 191-220.

2. Yau, N. (2013), pp. 165-176.

3. Tableau Online Help – Mapping at http://kb.tableausoftware.com/articles/knowledgebase/mapping-basics

1.1.2. Mapping Point-to-Point Details on Map

In Tableau, mapping point-to-point details on map requires that the data supports plotting and linking each points.

Using the e-commerce company scenario, when the Sales Vice President of the company would like to see the linking between account manager locations and customer locations, he can use Map and linking each points (i.e. the points represent account manager locations and customer locations) on the Map in Tableau as illustrated in Figure 5.2. Below are the steps to creating a Point-to-Point Map in Tableau:

1. Create a new worksheet.

2. Create two rows of data for each path we want to trace i.e. the “from and to”.

3. For each data pair, create a unique key that identifies them as a pair.

4. Create a column to define the Path Order, i.e. the direction in which the line is drawn.

5. Drag “Longitude” from the measure pane to the Columns shelf, and “Latitude” to the Rows shelf.

6. On the shelves, right click on the Longitude and the Latitude and select the Dimension.

7. Drag the “Unique Key” that identifies from and to information onto the “Level of Detail”.

8. On the Marks Card, the Marks type drop down list, select “Line”.

9. From the “Dimension” pane, drag “Path Order” to the Path shelf. If we want to see a path over time, we have to put the “Date” field on the Path shelf.

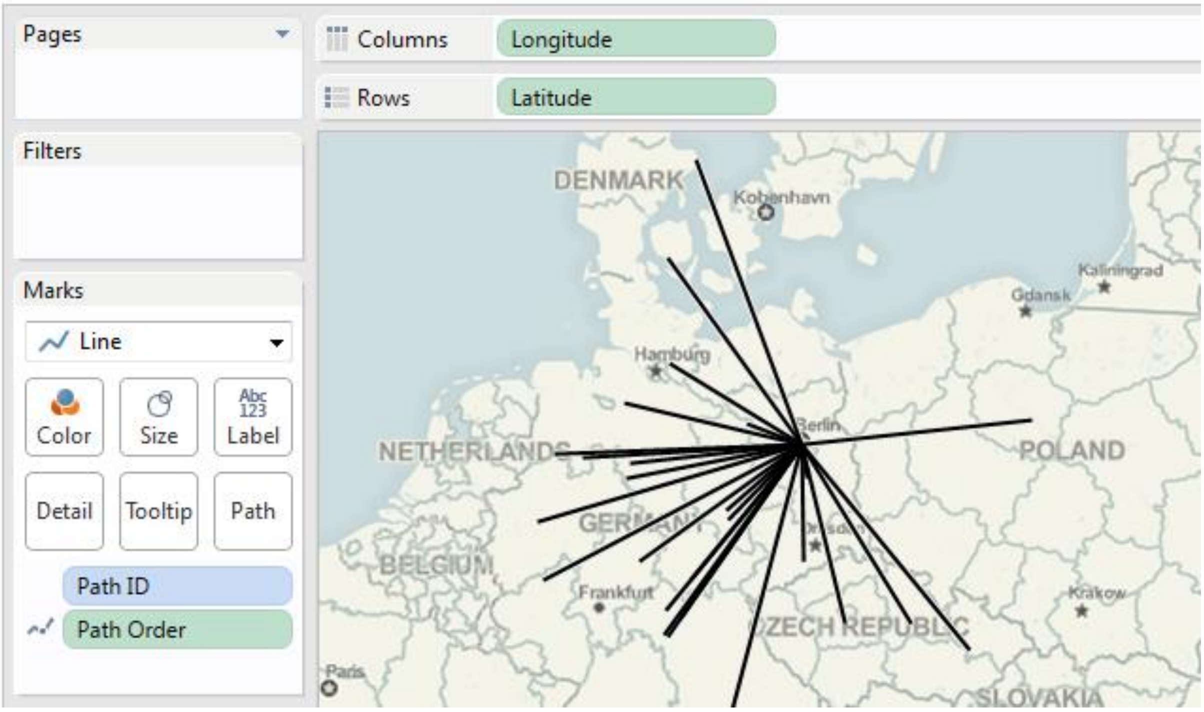


Figure 5.2: Mapping point-to-point detail

READ

Here is an optional reading to serve as an additional reference for students who want to learn more about the topics just covered:

1. Tableau Online Help – Using the Path Shelf for From – To Pattern Analysis at http://kb.tableausoftware.com/articles/knowledgebase/using-path-shelf-pattern-analysis.

2. Visualisation for Multi Variable Data

WATCH

Visualisation for Multi Variable Data
(Access video via iStudyGuide)

According to Yau (2013), data often come in a table format with multiple columns. Each column represents a variable. These data may come from survey questionnaires, results of an experiment that measure multiple aspects of a system, or demographic data on countries that include multiple information of each country.

2.1. Best Practices to Visualise Multi Variable Data

To visualise multi-variable data, we can fit all data onto a screen and display the relationships amongst variables or trends in each variables. However, sometimes, the relationships amongst variables or trends are not so straightforward. In such cases, we need to use multiple views through more straightforward charts.

We can use scatter plot to visualise the relationships amongst variables. We can plot one variable on the horizontal axis, and the other variable on the vertical axis. The statistical relationship amongst variables is called correlation. Correlation is a statistical measure that measures the degree of how one variable's value tends to change in a certain way as the other variable's value changes. For example, the price of taxi fare and the price of gasoline are positively correlated as both have been increasing over the years.

As we analyse the relationships amongst variables, we should not confuse between correlation and causation. Causation is the extent of how one variable will impact another. To differentiate between correlation and causation, using the taxi fare and gasoline prices as an example, we may need to ask whether the price of taxi fare goes up by default if we increase the price of gasoline. Next, if the price of taxi fare does go up, is it because of the increase in the price of gasoline or due to outside factors such as taxi rental. It is difficult to account for outside factors, thus it is challenging to prove the causation. It is however easy to find correlation. Correlation can help us predict one variable's value by knowing the other variable's value.

READ

Here is an optional reading to serve as an additional reference for students who want to learn more about the topics just covered:

1. Yau, N. (2013), pp. 176-189, 189-199

2.2. Creating Charts to Visualise Multi Variable Data

2.2.1. Scatter Plot, Circle View and Side-by-Side Circle Plots

To analyse granular data across multiple variables, scatter plots, circle views and side-by-side circles can be used to identify the outliers.

Scatter plot uses horizontal and vertical axes to compare two variables. It also uses colour, shape and size to express more aspects of the variables. Circle view uses horizontal and vertical axes to compare two variables. Side-by-Side circle plot displays more than one circle view side-by-side to provide granular breakdown of the circle view, based on a variable.

Using the e-commerce company scenario, when the Sales Vice President of the company would like to compare the profit and the order quantity, he can use Scatter Plot in Tableau as illustrated in Figure 5.3. Below are the steps to creating a Scatter Plot in Tableau:

1. Create a new worksheet.

2. Drag one measure into the worksheet's columns. The measure is automatically aggregated as a summation. An axis is created with a label given by the name of the field.

3. Drag one measure into the worksheet's rows. The measure is automatically aggregated as a summation. An axis is created with a label given by the name of the field.

4. Disaggregate the measures by right clicking on each measure, and select "Dimension".

5. Select "Aggregate Measures" from the analysis menu to display all the variables as a scatter plot.

6. The Tableau will automatically select "Mark as Shape" as an open circle and the mark colour as blue.

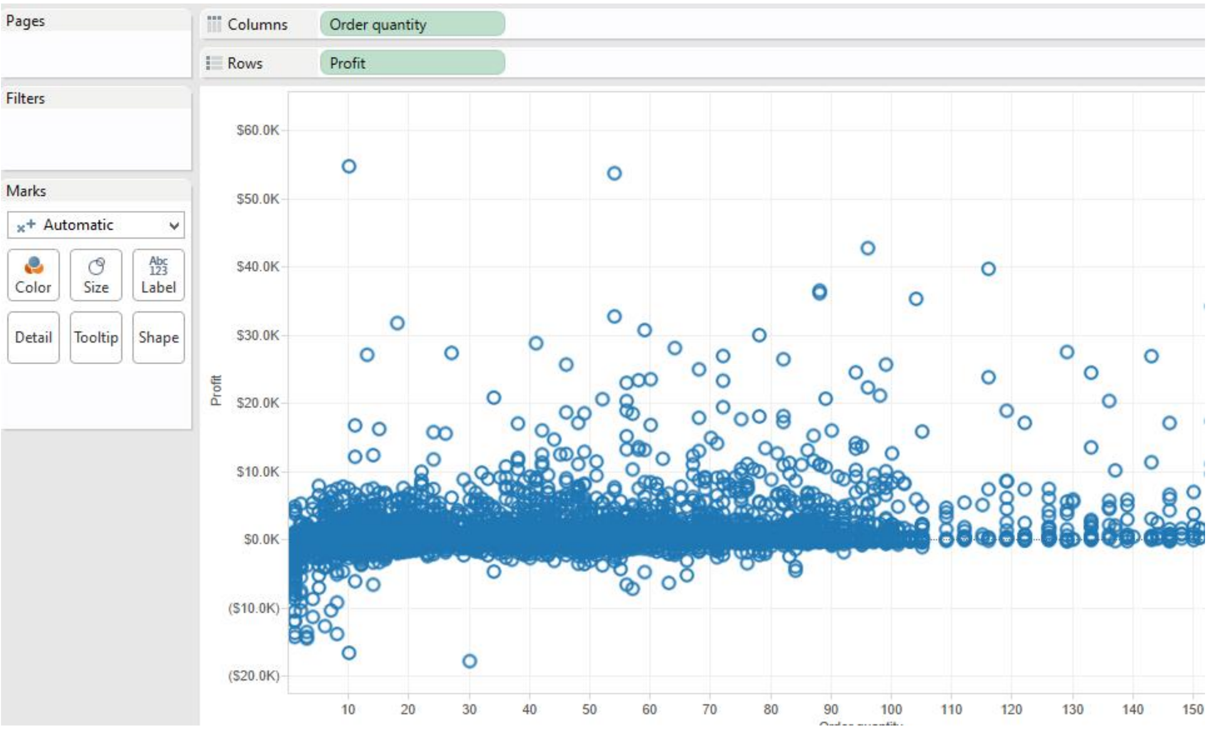


Figure 5.3: Scatter Plot

Using the e-commerce company scenario, when the Sales Vice President of the company would like to compare the sales of year 2013 in each region and department, he can use Circle View in Tableau as illustrated in Figure 5.4. Below are the steps to creating a Circle View in Tableau:

1. Create a new worksheet.

2. Drag one or more dimensions into the worksheet's columns.

3. Drag one measure into the worksheet's rows.

4. Select "Circle Views" as the chart type.

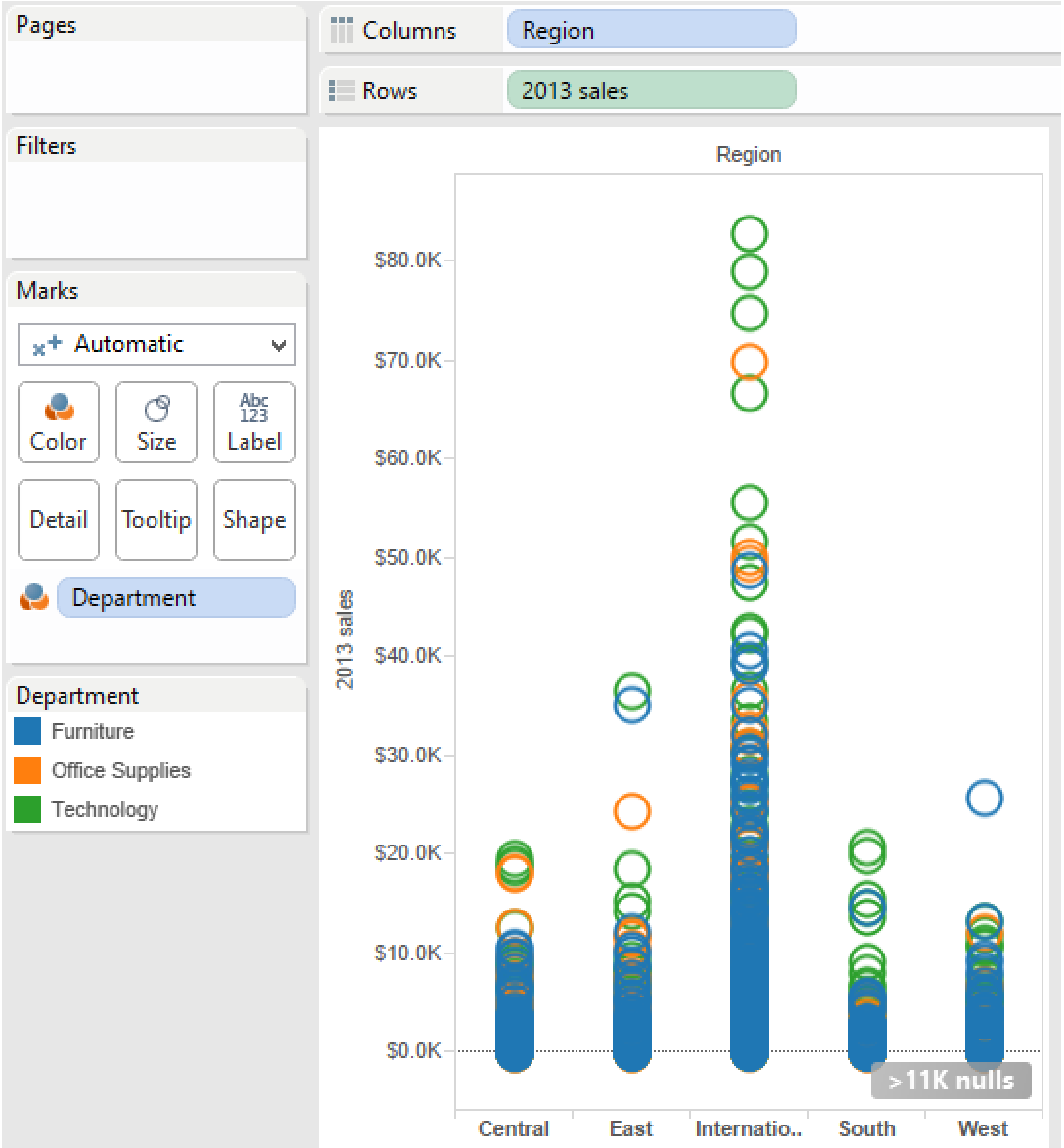


Figure 5.4: Circle View

Using the e-commerce company scenario, when the Sales Vice President of the company would like to compare the sales in year 2013 in each region and department. Instead of using Circle View, he can use Side-by-Side Circle View in Tableau as illustrated in Figure 5.5 to provide granular breakdown of the Circle View, based on region and department. Below are the steps to creating a Side-by-Side Circle Plot in Tableau:

1. Create a new worksheet.

2. Drag one or more dimensions into the worksheet's columns.

3. Drag one measure into the worksheet's rows.

4. Select "Side-by-Side Circles" as the chart type.

Figure 5.5: Side-by-Side Circle Plot

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Yau, N. (2013), pp. 176-189.

2. Tableau Online Help – Scatter Plot at http://onlinehelp.tableausoftware.com/v8.0/pro/online/en-us/buildexamples_scatter.html

3. Visualisation for Distribution of Data

WATCH

Visualisation for Distribution of Data
(Access video via iStudyGuide)

In addition to multi-variable data, we also often hear about mean and median. The former implies the middle or average value of a data within a group of people, places or things; and what the above or below average measurement is. However, we should also know more about the distributions of our data, rather than just where the middle or average value of our data is.

3.1. Best Practices to Visualise Distribution Data

We can visualise the distribution of our data at different granularities with bullet chart, bubble chart, and histogram. For distribution of one variable, using histogram will enable us to see where the data is clustered and any outliers, by keeping track of where the outliers sit on the vertical axis. For distribution of multiple variables, sometimes values come as pairs, so it makes more sense to visualise both values at the same time.

3.2. Creating Charts to Visualise Distribution of Data

3.2.1. Bullet Chart, Bubble Chart and Histogram

A Bullet Chart is a bar chart that includes reference lines and reference distributions for each cell in the chart.

A Bubble Chart display one-to-many comparison by using size and colour. However, it does not allow precise comparison between different bubbles. Therefore, we should use Bubble Chart when we do not require precise visual ranking of the bubble.

Histogram turns continuous variables into discretely bucketed bins of variable. It uses bar length to display the value of the variable.

Using the e-commerce company scenario, when the Sales Vice President of the company would like to visualise the sales in years 2012 and 2013 in each region together with target sales, he can use Bullet Chart in Tableau as illustrated in Figure 5.6. Below are the steps to creating a Bullet Chart in Tableau:

1. Create a new worksheet.
2. Drag one or more dimensions into the worksheet's rows.
3. Drag two measures into the worksheet's columns.
4. Choose "Bullet Chart" as the chart type.

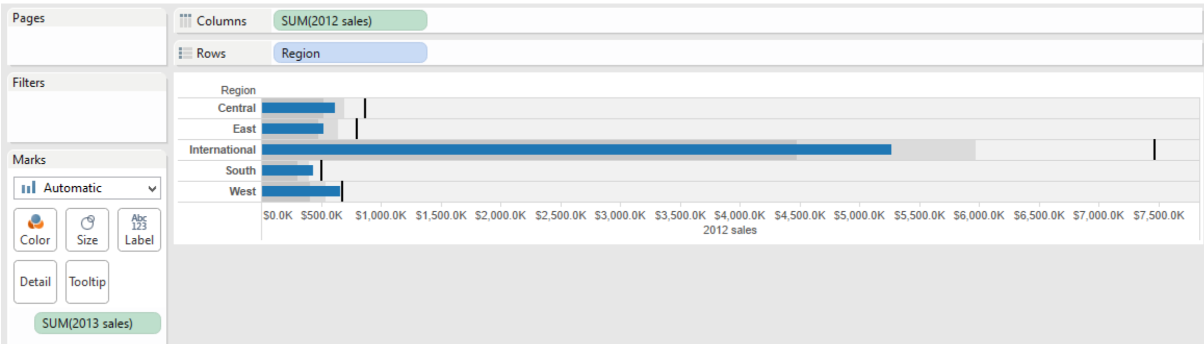


Figure 5.6: Bullet Chart

Using the e-commerce company scenario, when the Sales Vice President of the company would like to compare the profit and the sales in year 2013 in each region and department, he can use a Bubble Chart in Tableau as illustrated in Figure 5.7. Below are the steps to creating a Bubble Chart in Tableau:

1. Create a new worksheet.
2. Drag one or more dimensions (i.e. Department) into the worksheet's columns.
3. Drag one or more measures (i.e. Sales 2013) into the worksheet's rows.
4. Choose "Packed Bubbles" as the chart type.
5. Drag one dimension (i.e. region field) into the Marks Card to get more bubbles in view.
6. Drag one measure (i.e. profit field) into "Colour" on the Marks Card to colour code the bubbles based on the measure's value.
7. Drag the dimension in Step 5 into "Label" on the Marks Card to add region label to the bubble.

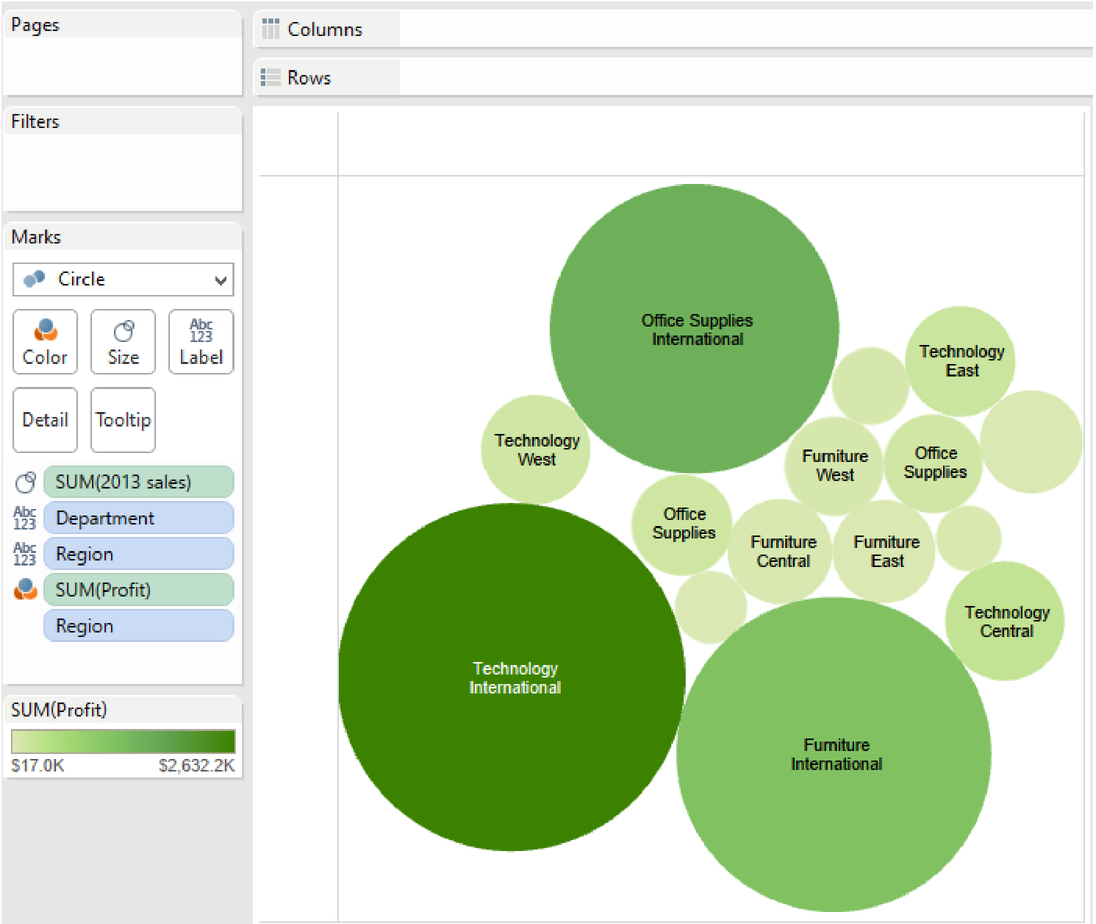


Figure 5.7: Bubble Chart

Using the e-commerce company scenario, when the Sales Vice President of the company would like to visualise the distribution of the sales in year 2013, he can use Histogram in Tableau as illustrated in Figure 5.8. Below are the steps to creating a Histogram in Tableau:

1. Create a new worksheet.

2. Right click on one measure (i.e. Sales 2013) in the “Data” window and select “Create Bins”.

3. On the “Create Bins” interface, key in “Sales Bin” under “New Filed Name”, and “1000” under “Size of Bins”. Click the “Load” button and it will generate Min, Max and Diff values. Finally, click on “OK” to close the “Create Bins” interface.

4. Drag one measure (i.e. Sales 2013) into the worksheet’s rows.

5. Drag the “Sales Bin” into the worksheet’s columns.

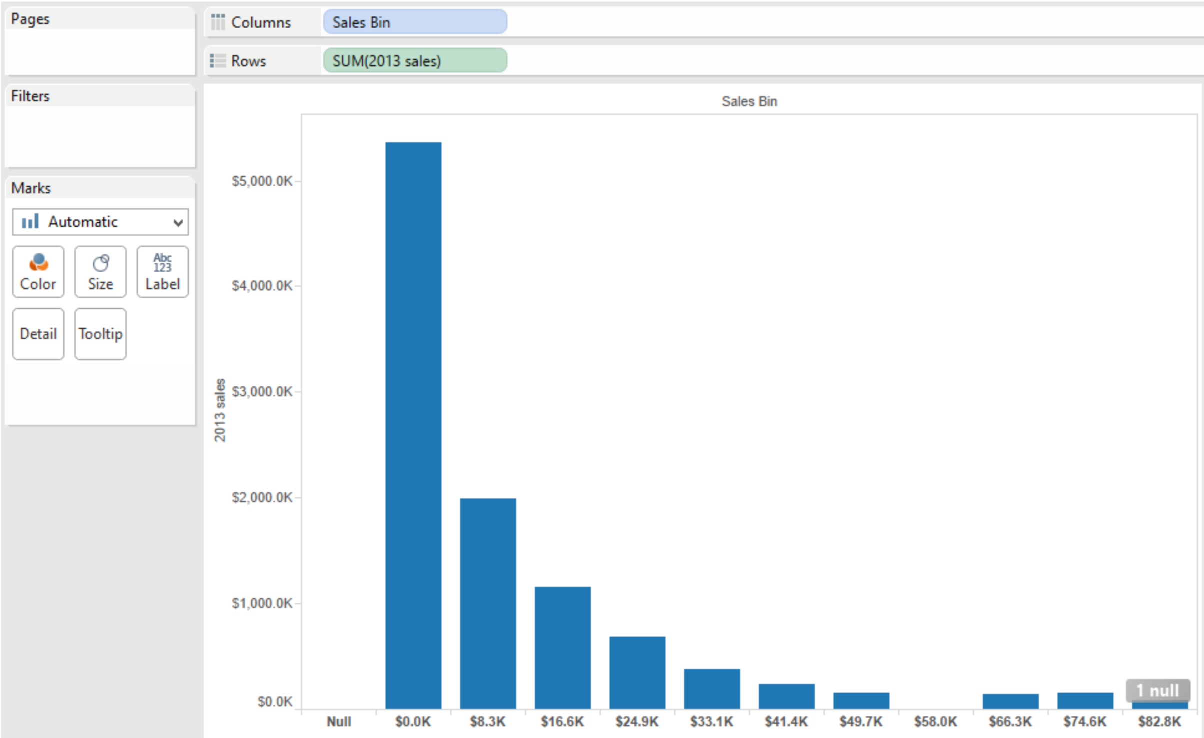


Figure 5.8: Histogram Chart

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Yau, N. (2013), pp. 149, 197-198,.
2. Tableau Online Help – Bullet Charts at http://kb.tableausoftware.com/articles/knowledgebase/bulletgraph.
3. Tableau Online Help – Bubble Charts at http://kb.tableausoftware.com/articles/knowledgebase/bulletgraph.
4. Tableau Online Help – Histogram at Tableau Online Help – Bubble Charts at http://kb.tableausoftware.com/articles/knowledgebase/bulletgraph.

3.2.2. Box Plot and Pareto Chart

By using a Box Plot, we are able to display fine granular distribution of a measure across multiple members of a dimensions set. Pareto Chart is known as 80-20 rule, developed by Vilfredo Pareto in 1960. It was developed to describe the unequal distribution of wealth in his country. In general, the 80-20 rule states that 20 percent of the inputs account for 80 percent of the outputs.

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Yau, N. (2013), pp. 195-196.

2. Tableau Online Help – Box Plot at http://kb.tableausoftware.com/articles/knowledgebase/box-plot-analog.

3. Tableau Online Help – Pareto Charts at http://kb.tableausoftware.com/articles/knowledgebase/pareto-analysis.

4. Forecasting

WATCH

Forecasting
(Access video via iStudyGuide)

Predicting future values is in nature imprecise. Using extensive data from the past will give us an idea of what to expect in the future. Forecasting is the act of predicting future values based on historical values.

4.1. Building Parameters

Parameters allow data visualisation to change the context of views with quick filter like controls. This will be useful for non-technical data visualisation users to conduct ad-hoc analysis by changing what, and how dimensions and measures are displayed, within the boundaries of the designer's intended usage.

Parameters are variables that allow data visualisation users to alter the content of a formula or change a dimension or measure contained in the view. Using parameters, data visualisation users can change normally static values into dynamic entities that facilitate ad-hoc analysis without the need to change the design of the data visualisation.

In Tableau, there are three basic parameter types that are built for typical-use cases that benefit from variables. They are:

- Reference Line parameters
- Bin Size parameters for histogram
- Ranking parameter for in value comparison view

READ

Here is an optional reading to serve as an additional reference for students who want to learn more about the topics just covered:

1. Tableau Online Help – Using Parameters at http://kb.tableausoftware.com/articles/knowledgebase/using-parameters.

4.2. Creating Forecast

In Tableau, we can forecast data in different ways. Tableau will recommend what it deems is the best method to forecast our data. If we do not want to accept the default method, we can always edit the forecast model by right-clicking on the worksheet and selecting the forecast option menu.

Tableau provides the following forecast trend models:

- Trend and season
- Trend only
- Season only
- No trend or season

Depending on the amount and granularity of data, each of the above forecast trend model generates different results. Trend and season model will generate the most volatile forecast data. The number of periods that Tableau forecasts is dependent on the data range and on the data aggregation level presented in the data visualisation.

READ

Here are some optional readings to serve as additional references for students who want to learn more about the topics just covered:

1. Murray. D. G. (2016), pp. 221-242.

2. Tableau Online Help – Forecasting at http://www.tableausoftware.com/new-features/forecasting.

Summary

In visualising spatial data, we map data to colour and geometry on a map. It seems easy because we just need to draw and colour the geometry. However, it is challenging to figure out what shapes and colours work best, where to put them, and how to size them.

Analysing relationships in our data can be challenging and it requires more critical thinking than blindly graphing measures and dimensions. But it can also be the most rewarding and informative method for us, because we will be able to gain an actionable insight by understanding the relationships in our data. It is how our data relate and interact among themselves. By knowing the relationships in our data, we will be able to tell the best stories behind the data.

Regardless of the type of data visualisation we use to explore the distribution of data, try to observer the peaks and valleys, range, and the spread of our data. They tell us a lot more than just the mean and median would. The visual analysis of raw data and the variation in between the summary statistics will help us derive insights from our data.

Tableau provides the best fit forecast that data visualisation users can modify by selecting from a menu of available forecast options, including automatic, automatic with seasonality, trend and season, trend only, season only, or no trend or season.

Exercises

ACTIVITY 1

Prepare your answers on below questions as your instructor will ask you to present your answers in front of the class:

1. Briefly explain the best practices to visualise spatial data with examples.
2. Briefly explain the best practices to visualise multi-variable data with examples.
3. Briefly explain the best practices to visualise distribution of data with examples.
4. Explain in what scenario we need to create forecasting in data visualisation.

ACTIVITY 2

1. Using the country data and Tableau workspace prepared in Week 2 and Week 4, create:

   a. Point-to-Point Map
   b. Scatter Plot, Circle View and Side-by-Side Circle Plots
   c. Bullet Chart, Bubble Chart and Histogram
   d. Box Plot and Pareto Chart
   e. Forecasting

2. The Instructor will randomly appoint students to present their work in front of the class.

Tableau Demos

Mapping

This demo aims to present how to create Map using Tableau Desktop.

At the end of this demo, you will be able to:
- create Symbol and Filled Map
- create Dual Axis Map
- edit locations

Click here to view the demo.

Histogram

This demo aims to present how to create Histogram using Tableau Desktop.

At the end of this demo, students will be able to:
- create Histogram
- create Cummulative Histogram

Click here to view the demo.

Bubble Chart

This demo aims to present how to create Bubble Chart using Tableau Desktop.

At the end of this demo, students will be able to:
- create Bubble Chart

Click here to view the demo.

Box Plot

This demo aims to present how to create Box Plot using Tableau Desktop.

At the end of this demo, students will be able to:
- create Box Plot

Click here to view the demo.

Parameters

This demo aims to present how to create Parameters using Tableau Desktop.

At the end of this demo, students will be able to:
- create Parameters
- swap dimensions with Parameters
- perform "what if" analysis using Parameters

Click here to view the demo.

Forecasting

This demo aims to present how to create Forecasting using Tableau Desktop.

At the end of this demo, students will be able to:
- create Trend Lines
- create Forecasting

Click here to view the demo.

References

Murray, G. D. (2016). *Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software. (2nd Edition)* Indianapolis, IN, USA: John Wiley & Sons Inc.

Tableau Online Help. (2014). Retrieved on 1 July 2014 from http://onlinehelp.tableausoftware.com/current/pro/online/mac/en-us/help.htm#default.html.

Yau, N. (2013). *Data Points: Visualisation that Means Something*. Indianapolis, IN, USA: John Wiley & Sons.

STUDY UNIT 6
BUSINESS PERFORMANCE DASHBOARD

Learning Outcomes

At the end of this unit, you are expected to:

- explain the concepts and benefits of a Business Performance Dashboard
- compare Strategic, Tactical and Operational Dashboards
- explain the wrong and right ways in building a dashboard
- illustrate dashboard design principles
- arrange the objects in the dashboard workspace
- use actions to create advanced dashboard navigation

Overview

This unit introduces the concepts of business performance dashboard. In addition, it also discuss techniques in building a business performance dashboard.

1. Introduction to Business Performance Dashboard

WATCH

Introduction to Business Performance Dashboard
(Access video via iStudyGuide)

A *Business Performance Dashboard* provides useful information that enables managers to mount improvement activities on areas that show performance below expectation. Lastly, the Business Performance Dashboard provides information on an organisation's attainment level of strategic goals at a specific point in time.

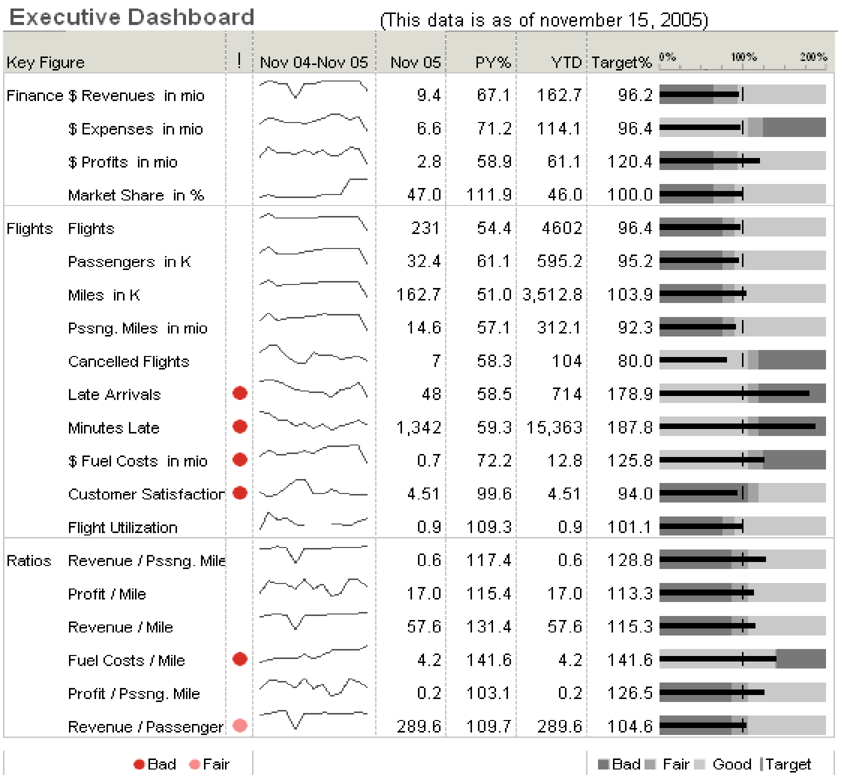
A Business Performance Dashboard provides concise and easily understandable visual representations of performance in many critical business processes. This allows the management to know the business processes that are doing well and those that are having problems and requiring corrective actions. Ultimately, the Business Performance Dashboard serves as a critical performance reporting function in the strategic performance management model. It acts like a magnifying glass to bring to focus an organisation's attention on deficiencies that impede the achievement of the overall strategy.

Eckerson (2006) found strong interest in Business Performance Dashboards in a survey among IT professionals. At least 43 percent of the 689 respondents implemented or are in the process of implementing Business Performance Dashboards. According to the profile of the respondents, this interest is not only limited to those in the manufacturing industry, but also cuts across all industries, including the government sector. The basic functionalities of a Business Performance Dashboard according to Eckerson are to

- monitor critical business processes and activities, using the relevant measures that track business performance; and

- manage people and processes to improve decisions making, optimise performance, and steer the organisation in the same direction.

Figure 6.1 illustrates an award-winning Business Performance Dashboard of an airline company to track the company's relevant performance (Few, 2006).



Executive Dashboard (This data is as of november 15, 2005)

Key Figure		!	Nov 04-Nov 05	Nov 05	PY%	YTD	Target%	0% 100% 200%
Finance	$ Revenues in mio			9.4	67.1	162.7	96.2	
	$ Expenses in mio			6.6	71.2	114.1	96.4	
	$ Profits in mio			2.8	58.9	61.1	120.4	
	Market Share in %			47.0	111.9	46.0	100.0	
Flights	Flights			231	54.4	4602	96.4	
	Passengers in K			32.4	61.1	595.2	95.2	
	Miles in K			162.7	51.0	3,512.8	103.9	
	Pssng. Miles in mio			14.6	57.1	312.1	92.3	
	Cancelled Flights			7	58.3	104	80.0	
	Late Arrivals	●		48	58.5	714	178.9	
	Minutes Late	●		1,342	59.3	15,363	187.8	
	$ Fuel Costs in mio	●		0.7	72.2	12.8	125.8	
	Customer Satisfactior	●		4.51	99.6	4.51	94.0	
	Flight Utilization			0.9	109.3	0.9	101.1	
Ratios	Revenue / Pssng. Mile			0.6	117.4	0.6	128.8	
	Profit / Mile			17.0	115.4	17.0	113.3	
	Revenue / Mile			57.6	131.4	57.6	115.3	
	Fuel Costs / Mile	●		4.2	141.6	4.2	141.6	
	Profit / Pssng. Mile			0.2	103.1	0.2	126.5	
	Revenue / Passenger	●		289.6	109.7	289.6	104.6	

●Bad ●Fair ■Bad ■ Fair ■ Good |Target

Figure 6.1: Business Performance Dashboard for an airline company

Business Performance Dashboards can generally be classified into three main types. Each type of dashboard caters to a different group of users. The three types of dashboard are Strategic Dashboard, Tactical Dashboard, and Operational Dashboard.

1.1. Strategic Dashboard

The *Strategic Dashboard* is used by executives and managers to monitor the execution of strategic objectives. These strategic objectives can be the results of using a Balanced Scorecard approach. The key features of the Strategic Dashboard are simplicity and high visual impact. An intuitive Strategic Dashboard is something that allows the top management who often have a tight schedule to have a quick and fast look of the company's performance. If required, the top management is able to carry out further analysis by drilling down to the details of the respective object or to the underlying tactical dashboard.

Figure 6.2 illustrates an example of a Strategic Dashboard.

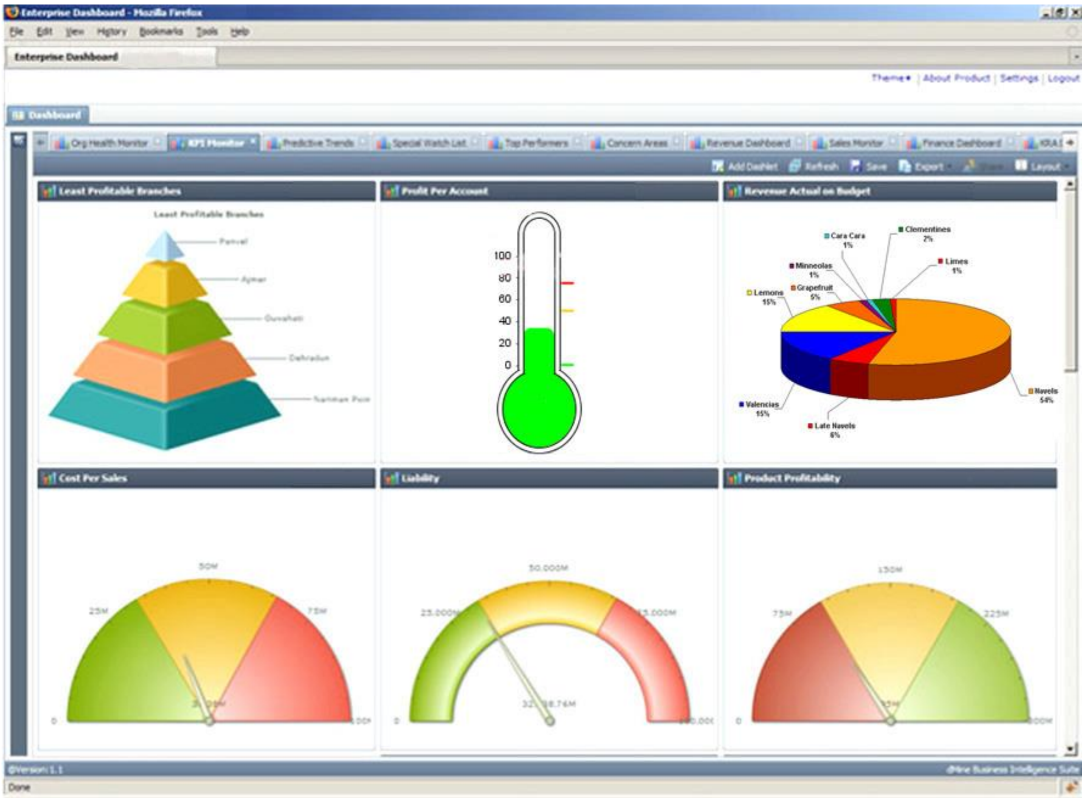


Figure 6.2: Strategic Dashboard

1.2. Tactical Dashboard

A *Tactical Dashboard* is used by managers and analysts to track the progress of departmental processes, and projects against budget plans, forecasts or last period's achievement. A Financial Dashboard is an example of a Tactical Dashboard used by the Finance Department to track revenue and expenses versus the budgeted numbers on a monthly or quarterly basis. A typical user group of Tactical Dashboard is business and financial analysts where they use multi-dimensional charts, slice-and-dice of Online Analytical Processing cube, etc.

Figure 6.3 illustrates an example of a Tactical Dashboard.

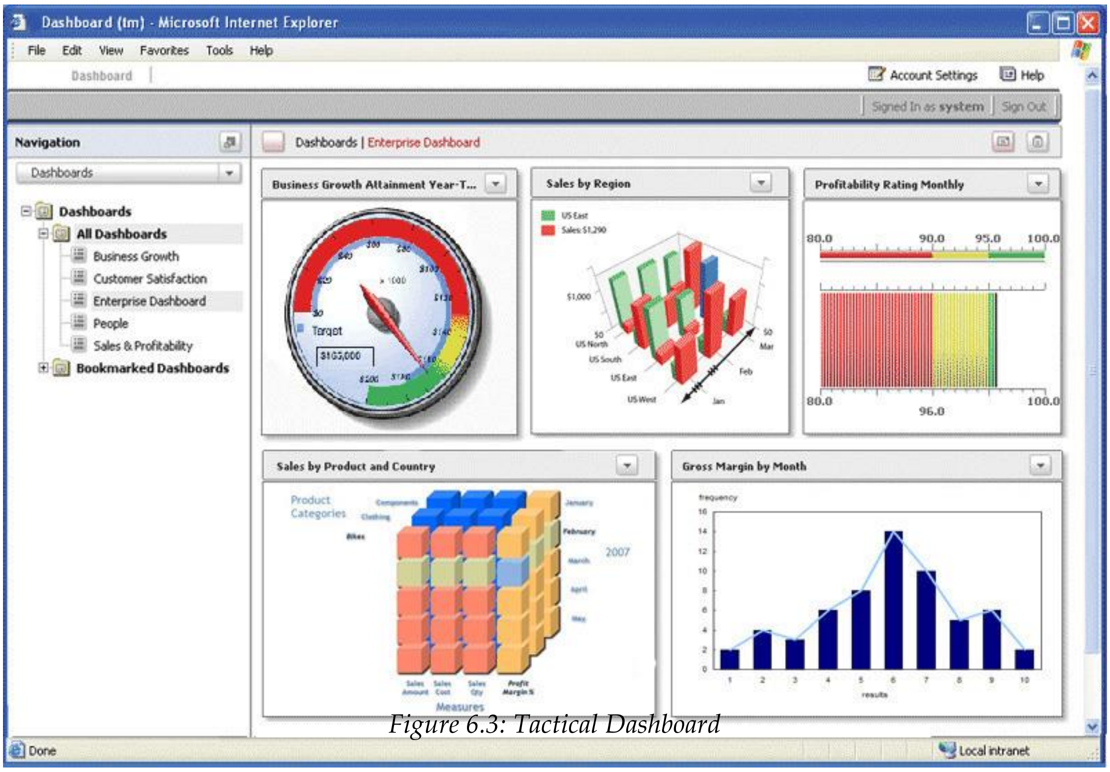


Figure 6.3: Tactical Dashboard

1.3. Operational Dashboard

The *Operational Dashboard* is used by operational staff and their supervisors to monitor operational processes. In a Sales Department, the Operational Dashboard is used to monitor the performance of individual sales representative in terms of sales achievement for a specific period of time, which is usually monthly or quarterly. In a manufacturing industry, the Operational Dashboard is used to monitor the performance of manufacturing engineers and assistants in terms of units of output for a day. Such dashboards are very specific to a particular operation or process.

Figure 6.4 illustrates an example of the Operational Dashboard.

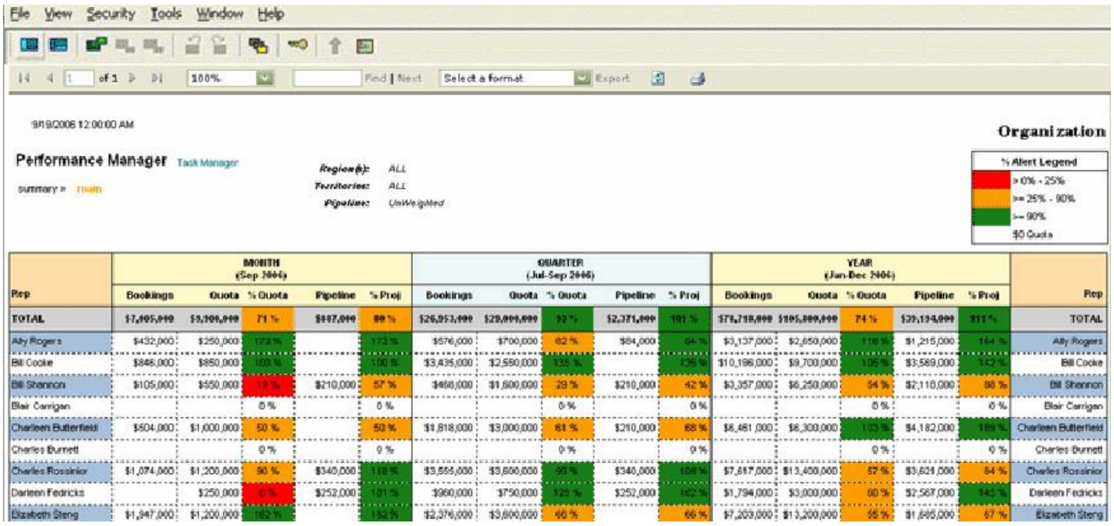


Figure 6.4: Operational Dashboard

1.4. Benefits of Business Performance Dashboard

Organisations implement Business Performance Dashboards because of the many benefits that can be reaped by everyone in an organisation. These benefits according to Eckerson (2006) are as follow.

1.4.1. Communicate Strategy

Business Performance Dashboards facilitate the translation of corporate strategy into measures with their corresponding targets. Anyone who has access to the Business Performance Dashboard will be aware of the organisation's strategic objectives and what needs to be done in his/her area to achieve these objectives.

1.4.2. Increase Insight

A Business Performance Dashboard empowers the management team with greater insights into the business performance in a timely manner. This minimises the element of surprise by unforeseen problems that affect the bottom-line or results. Corporate strategy can then be promptly fine-tuned with minor corrections, when required, instead of veering off-course.

1.4.3. Increase Motivation

With the measures and the corresponding targets publicly displayed in Business Performance Dashboards, the motivation to excel in the measured areas is increased.

Such motivation compels people to work harder out of pride, and the desire to achieve higher pay; especially when the pay moves in tandem with the measures. Knowing "what gets measured, gets done," transparency in performance measurement and results increases motivation to excel in the measured areas.

1.4.4. Increase Coordination

Objectivity, openness and transparency in Business Performance Dashboard help promote effective coordination, and encourage different departments to work more closely. Employees will be encouraged to engage in meaningful dialogue on performance results and forecasts. This allows managers to conduct more frequent constructive performance reviews.

1.4.5. Consistent View of Business

Business Performance Dashboard consolidates and integrates business performance information using common platform, definition, measures, and rules. A single version of such information minimises conflicts and disagreements.

1.4.6. Reduce Cost and Redundancy

Standardising information and reporting based on Business Performance Dashboard can eliminate redundant silo of informant and duplication of reports. A well designed Business Performance Dashboard is capable of replacing independent reporting systems, spreadsheet marts, data marts, and warehouses. This also helps streamline reporting and simplifies information system.

1.4.7. User Sufficiency and Empowerment

A Business Performance Dashboard allows users to be self-sufficient in creating, organising, and presenting information on business performance. Users are able to create the necessary materials promptly and efficiently without relying on the IT department or report writing specialist. With timely information, the management is able to resolve business issues promptly without wasting excessive time looking for the right information or report.

1.5. Attributes of Business Performance Dashboard

A Business Performance Dashboard can be defined as a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so that the information can be monitored at a glance (Few, 2004).

The original idea of the Business Performance Dashboard is taken from a vehicle dashboard, which provides all the measures that are critical to the operation of the vehicle, such as temperature and speed. Similarly, a Business Performance Dashboard serves to provide all the pertinent measures needed to drive the performance of an organisation.

The attributes of a Business Performance Dashboard are:

- Display of selected measures and their corresponding targets that track business performance. If the raw data is available, then the measures are usually high-level summaries of the raw data.

- Graphics that facilitate visualisation of past trends (using line plots) or magnitude of difference (using horizontal bar graphs).

- The use of traffic lights (green, ember, and red) to demarcate regions of desirability for a measure, as used in a temperature gauge shown in Figure 6.5.

Figure 6.5: Temperature Gauge – a traditional vehicle dashboard

In practice, such graphics may not be available in the Business Performance Dashboard, but the principle of traffic lights can be applied on the font colour of the measures as shown in the first column of the Airline Industry Dashboard shown in Figure 6.1. Another application of the traffic lights approach could be on directional icons that are often included next to the measures, as shown in the second column of Figure 6.6.

Measure with Traffic Lights (assuming target is 100)	**Corresponding Traffic Lights Directional Icon**
110	
90	
100	

Figure 6.6: Directional Traffic Light

2. Dashboard Design Principles

Dashboard Design Principles
(Access video via iStudyGuide)

The basic dashboard design principles taken from Alexander (2008) can be summarised as follow.

2.1. Keep It Simple

The well-known *KISS (Keep It Simple, Stupid)* principle of having simplicity as a key goal applies to Business Performance Dashboards design. It is important not to have too many measures and visual distractions so that attention is focused on the essentials.

2.2. Don't Display Everything

Fight the tendency to display everything, including raw data and all summaries of the data that one can think of. Display only what is required.

2.3. Keep to a Single Page

Eliminate the need to scroll back and forth, up and down to look at the measures. Since the dashboard should provide all relevant measures at a glance, segregate raw data from the actual dashboard whenever possible.

2.4. Avoid Fancy Formatting

Avoid using colours and other fancy formatting like pattern fills, gradients, shadows and glows on the dashboards. In fact, plain formatting is more effective in drawing attention to the measures, graphics, and the traffic lights.

2.5. Use Layout and Placement

Even after selecting the critical performance measures or components that need to be displayed on the dashboard, there may be differing levels of importance among them. This is where placement of these measures of a component can vary according to the level of importance. Alexander (2008) referred to studies done where the upper left and middle-left (i.e. dark blue) portion of a dashboard or document is called the Priority Zone, because it draws the most attention from dashboard viewers as shown in Figure 6.7. Hence, place the more important measures or components in the priority region of the dashboard.

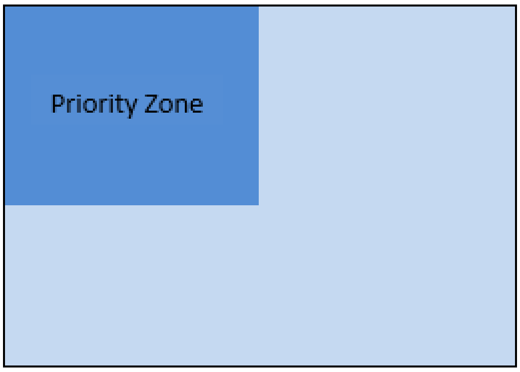


Figure 6.7: Priority zone in a dashboard

2.6. Format Numbers Effectively

Format numbers to display commas and dollar symbols when the measures are monetary values. Limit the number of decimal places (i.e. usually to two for currency values) and display in denominations of thousands or millions instead of the actual amount where appropriate (e.g. display 1.4 mil instead of 1,437,387).

2.7. Use Titles and Labels Effectively

Always remember to have descriptive titles, column headers, and footnotes, with minimum use of short forms and acronyms where possible. Include a time-stamp of when the report is generated or when data are extracted, where applicable.

3. Creating Business Performance Dashboard

Creating Business Performance Dashboard
(Access video via iStudyGuide)

A Business Performance Dashboard that relies on grids or overly complex charts does not communicate insights effectively to the viewers. It may take a longer time to visualise the data than the one that does not have complex charts and grids.

For example, a sales report displaying 12 months of history for 30 products (i.e. 12 x 30 = 360 data points) in a table, does not help dashboard viewers see the trends and outliers as easily as a time-series chart of the same information. Also, it will take a longer time to load the table than a time-series chart. Worst yet, if viewers need to use the web browser to view the dashboard and the Internet connection is slow, the interactivity of the table will be slower than the interactivity of the time-series chart.

The dashboard shown in Figure 6.8 displays another example of the wrong ways in building a dashboard. A pie chart has too many slices making it difficult for dashboard viewers to perform precise comparisons of each product (i.e. A, B, C, etc.).

A bar chart represents the sales margin and it displays on the y-axis ranges from zero to 90 percent, even though every data points are above 80 percent. It will be better to change the range of the y-axis so that it will give more zooming effect.

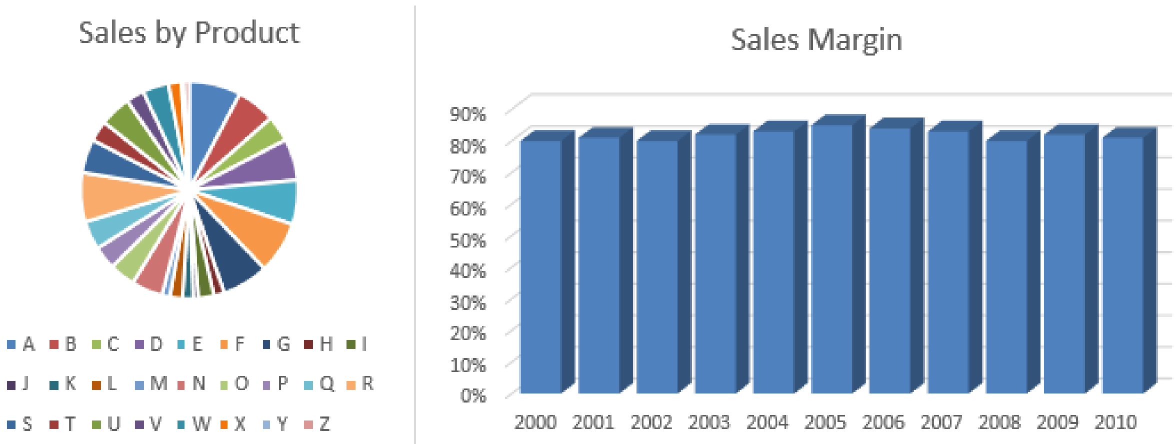


Figure 6.8: A wrong way in building a dashboard

3.1. Position Objects in the Dashboard Workspace

In a Tableau, after creating multiple worksheets, you can combine the worksheets to give an integrated view. Figure 6.9 shows an empty dashboard worksheet.

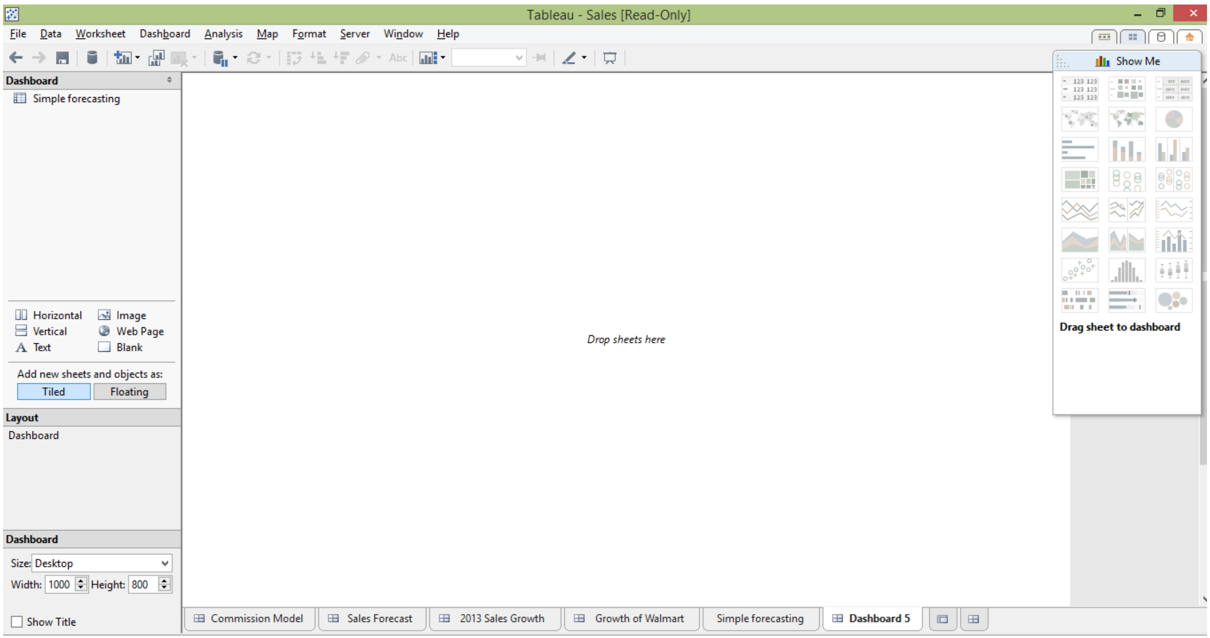


Figure 6.9: Tableau's dashboard worksheet

One of the first things we should consider when assembling worksheets in a dashboard is the available space that our viewers have to view. For example, whether the dashboard will be viewed on an overhead projector with limited resolution, or display the dashboard on a personal computer or smart devices.

The top-left hand corner of the dashboard worksheet displays all of the worksheets contained in the workbook. The bottom-left hand corner of the dashboard worksheet provides access to other object controls for adding text, images, blank space or live web pages into the dashboard worksheet. It also has control over specifying the size of the dashboard, and a checkbox for adding a dashboard title.

We can define the sizing of the entire dashboard and the individual objects included in the dashboard workspace. Before any worksheet is added into the dashboard workspace, we need to click the "size shelf", as shown in Figure 6.10, to define the size of the dashboard to accommodate the worst-case scenario in which the dashboard will be viewed. The automatic option expands the dashboard to fill up the available screen space. The exact option allows us to lock the dashboard width and height. The range option enables us to define minimum and maximum limits of the size of the dashboard. Once the size has been defined, we are ready to add individual worksheet objects into the dashboard.

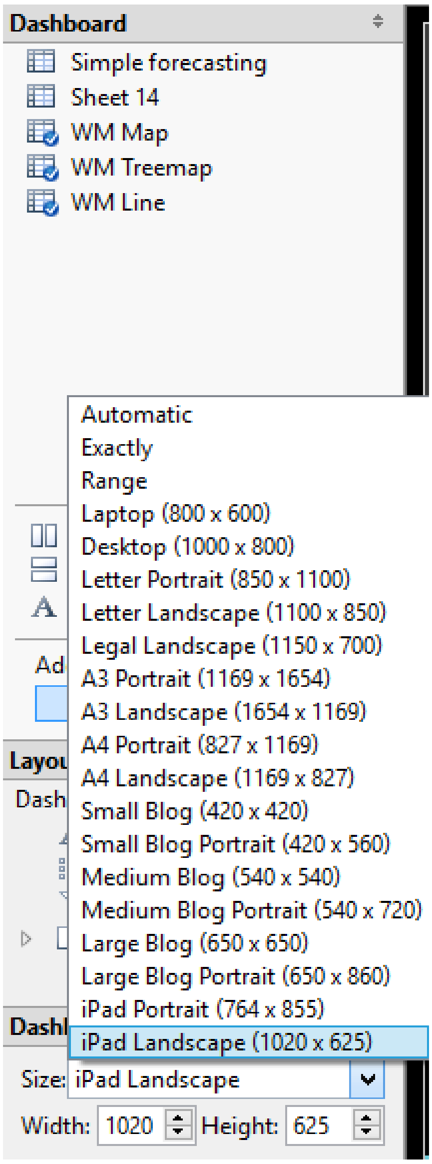


Figure 6.10: Tableau's dashboard layout size definition

There are two ways to add objects into the dashboard workspace. We can either drag the selected object into the "Drop sheet here" area, or double click the worksheet objects on the top-left hand corner. By double clicking on each worksheet objects, the Tableau will place that object onto the dashboard workspace automatically. To control the placement of an individual object more precisely, drag the object into the view. Once we lift up the left-click button, the Tableau will provide a preview of the area that the object will occupy by shading it in grey.

Subsequently, we need to add a title to the dashboard by selecting the "Show Title" option on the bottom left of the dashboard shelves. The default title will be the name of the dashboard worksheet. Edit the title text by double clicking on the default name and type in the new title. You can also edit the font size, colour, type and position of the title.

To make a dashboard communicates the information more effectively to its viewers, we need to follow these steps:

1. Ensure each worksheet object fits the entire view.

   To change the fit within the object pane, we can access the fit menu by right clicking on the title block of the object pane, or by right clicking on the title of the object pane. Figures 6.11 and 6.12 show how to access the fit menu using either method.

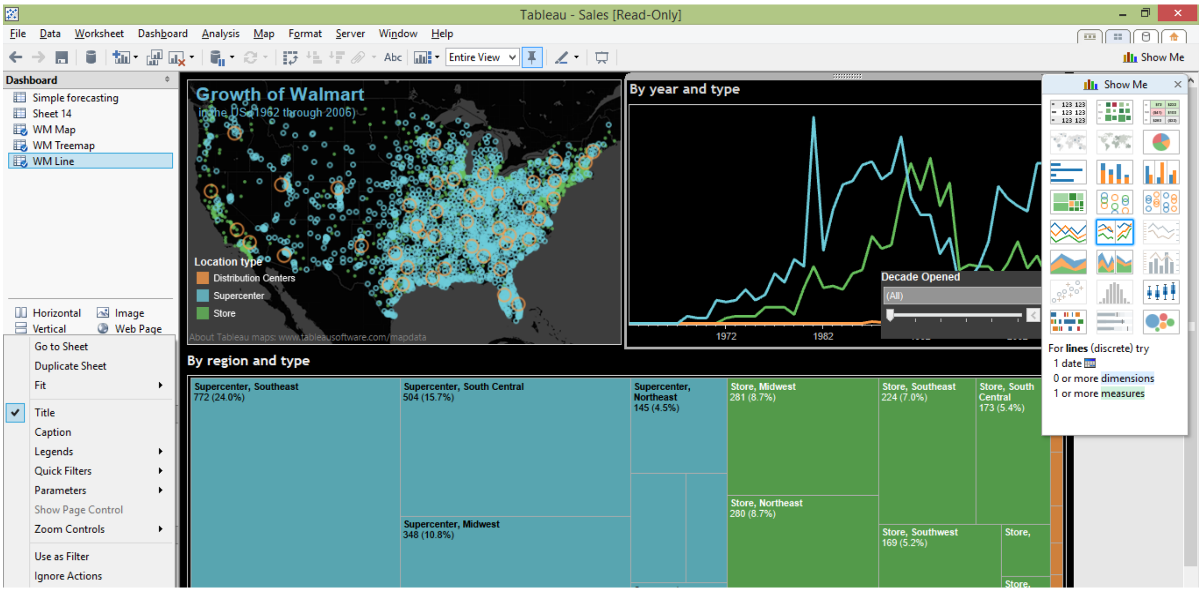


Figure 6.11: Accessing fit menu by right clicking on the title block of the object pane

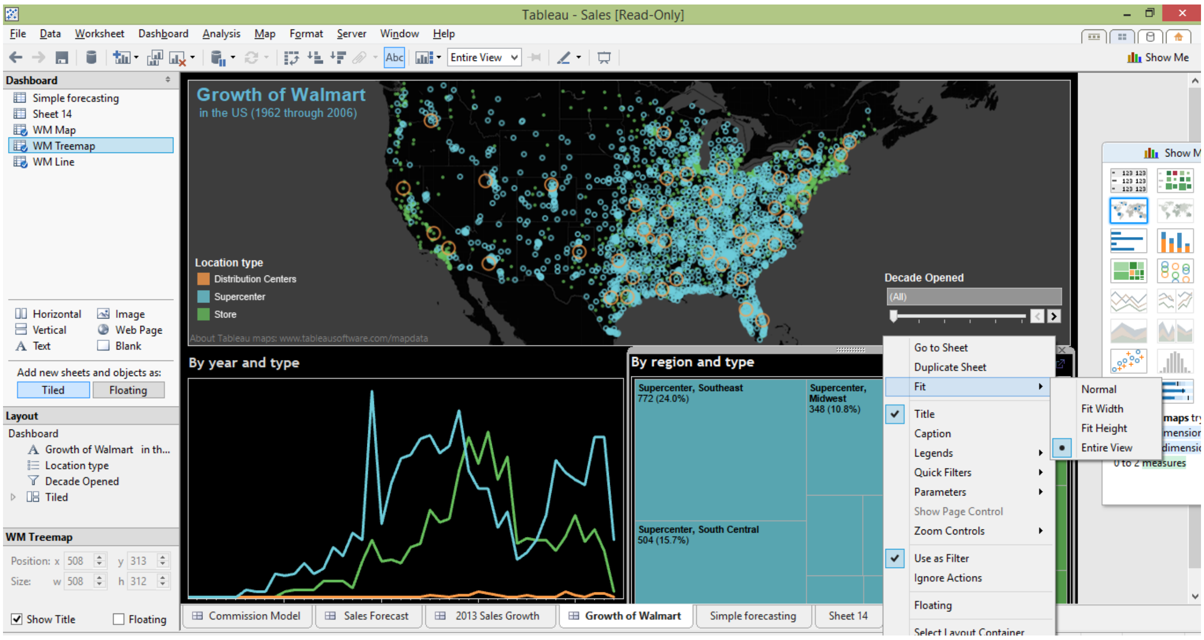


Figure 6.12: Accessing fit menu by right clicking on the title of the object pane

2. Create more descriptive titles for each chart.

   Adding more descriptive object titles will make it easier for viewers to interpret it. We can edit titles by double-clicking the object's title bar, as shown in Figure 6.13.

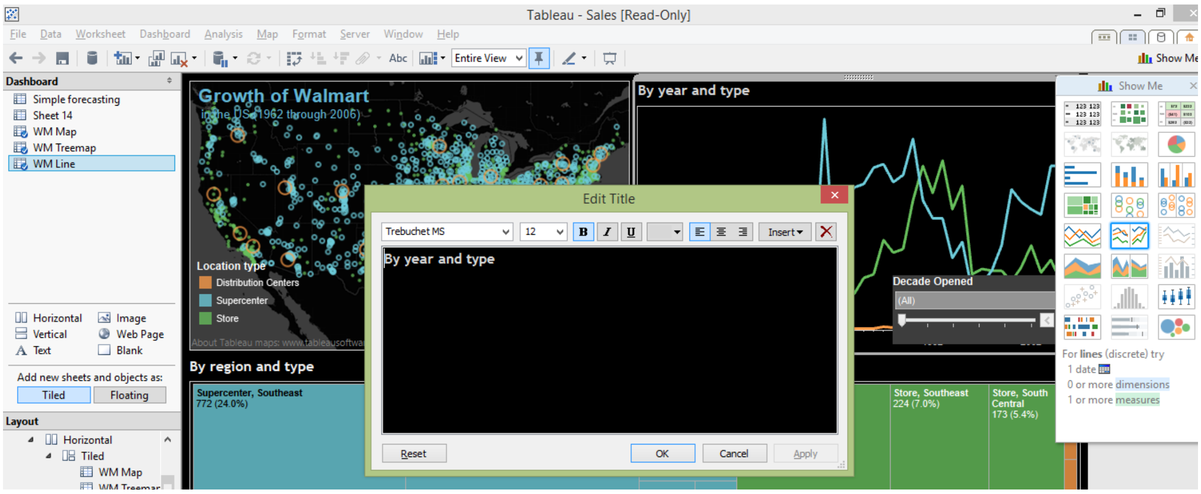


Figure 6.13: Accessing title edit menu by double-clicking the object's title bar

3. Make the attribute sort of charts in the same order.

   We can access the "sort menu" of each attribute by right-clicking on the field pill on the row shelf, then select the "sort menu" option as illustrated in Figure 6.14.

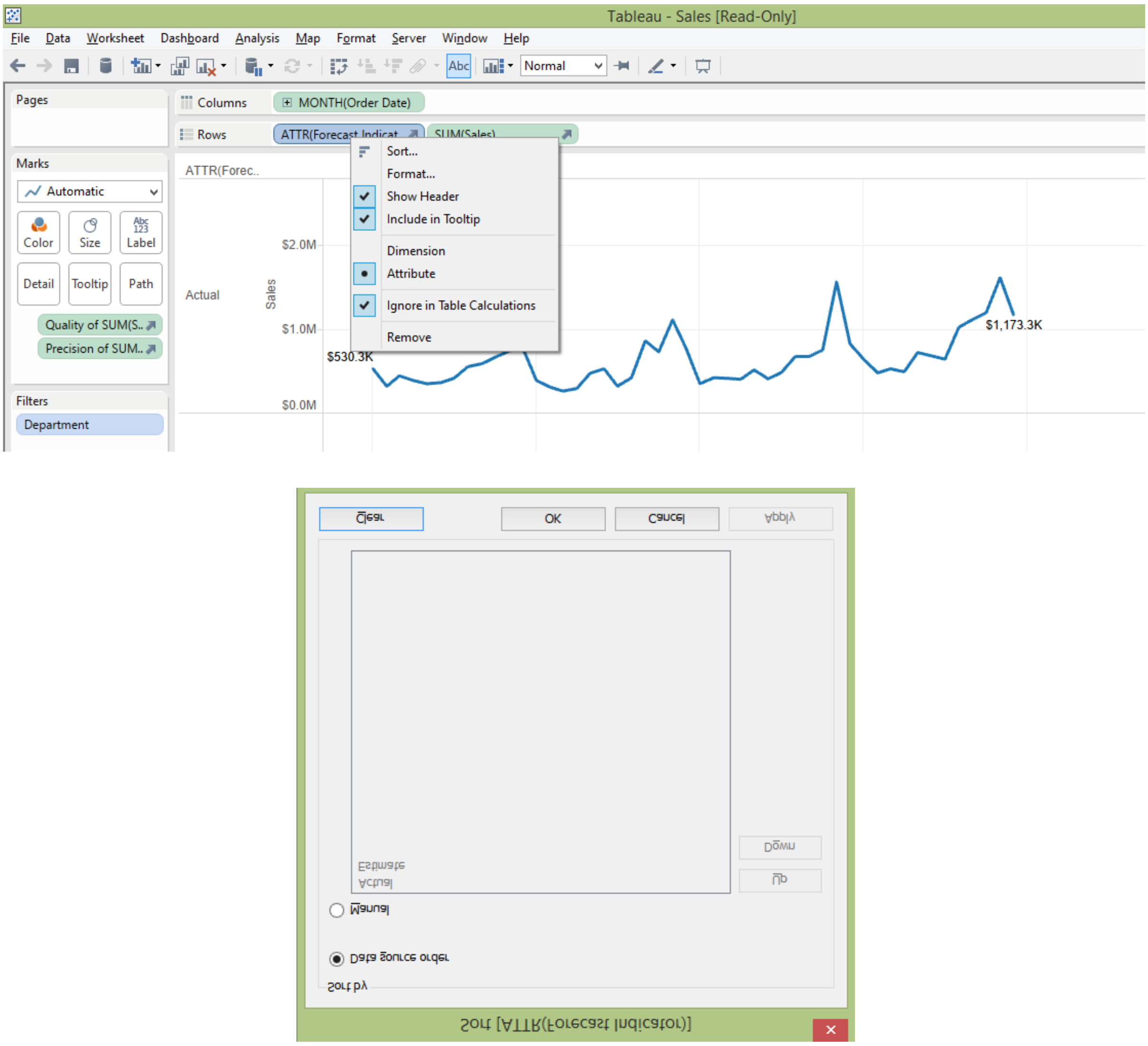


Figure 6.14: Accessing sort menu by right clicking on the field pill on the row shelf

4. Hide the axis headers and turn on "mark labels" in each chart.

   We can hide the axis headers in each chart by pointing at the axis header area, right-clicking and unchecking the "show header" option as shown in Figure 6.15. This will provide more space.

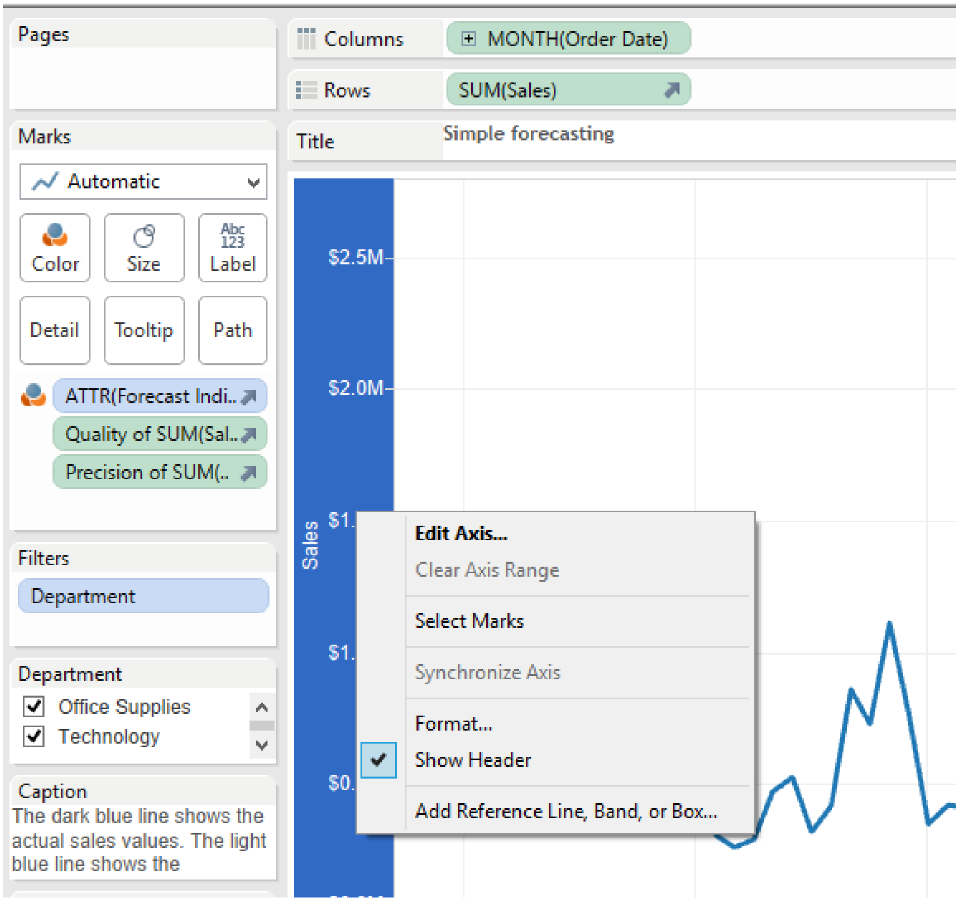


Figure 6.15: Hiding axis headers

It is good to hide the axis headers even though they provide valuable context. If the dashboard is going to be printed and consumed on paper, it is not a good idea to remove the axis headers.

When the dashboard is consumed interactively on a computer, we can use "mark labels" to replace axis headers by presenting important details on demand, when a mark is selected. Mark labels can always be displayed, but in this case space would be better utilised if mark labels are displayed only when the dashboard viewers want to see the labels. To make mark labels appear on demand, we can click the "label button" on the marks card to expose the label menu, as shown in Figure 6.16.

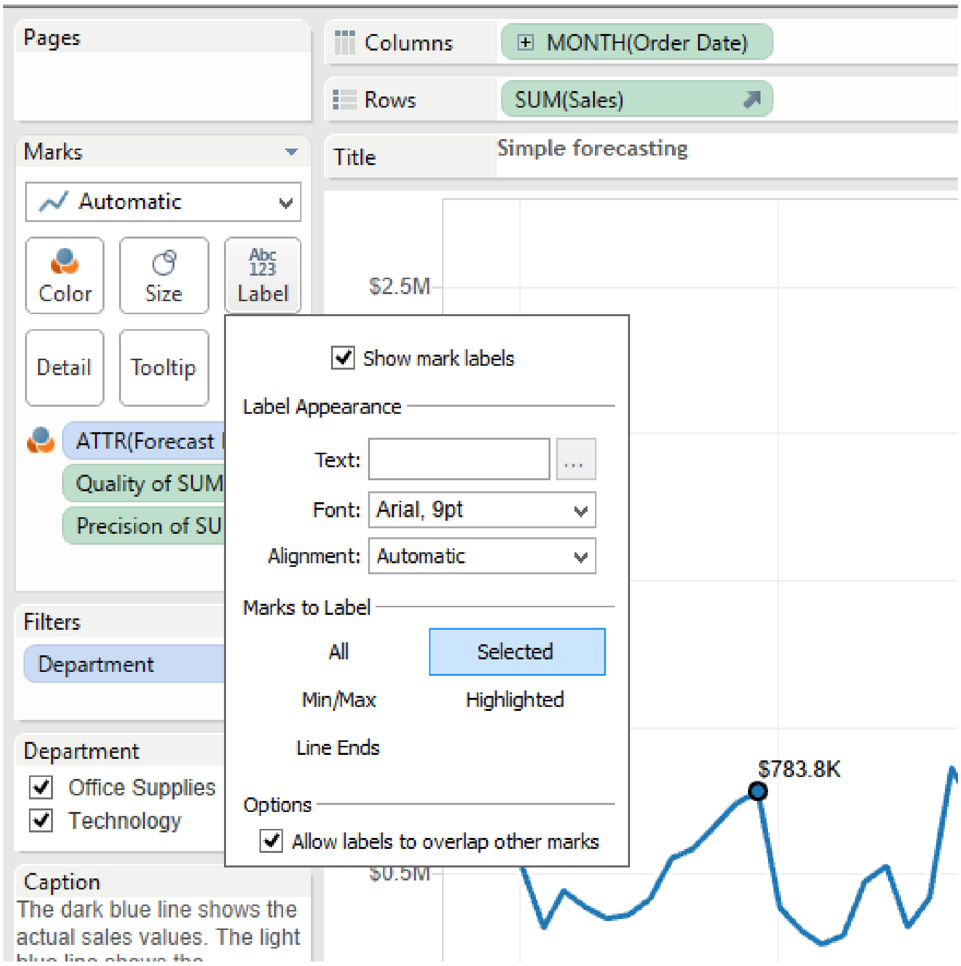


Figure 6.16: Showing mark labels

5. Reposition the worksheet objects to better use of space.
   When the chart legend causes a misalignment of the chart, we can place a text object to addresses the alignment needed and describe exactly what is being plotted on the chart.

3.2. Use Action to Create Advanced Dashboard Navigation

Actions facilitate insight discovery by altering the context of the dashboard based on selections made by the dashboard viewers. We can build actions that filter and highlight the main dashboard.

To create a filter action, we need to activate the object we want to filter, select the drop down arrow to expose the filter menu. Pick the "use as filter" menu option to create the filter action, and edit the filter action. Figure 6.17 illustrates these steps.

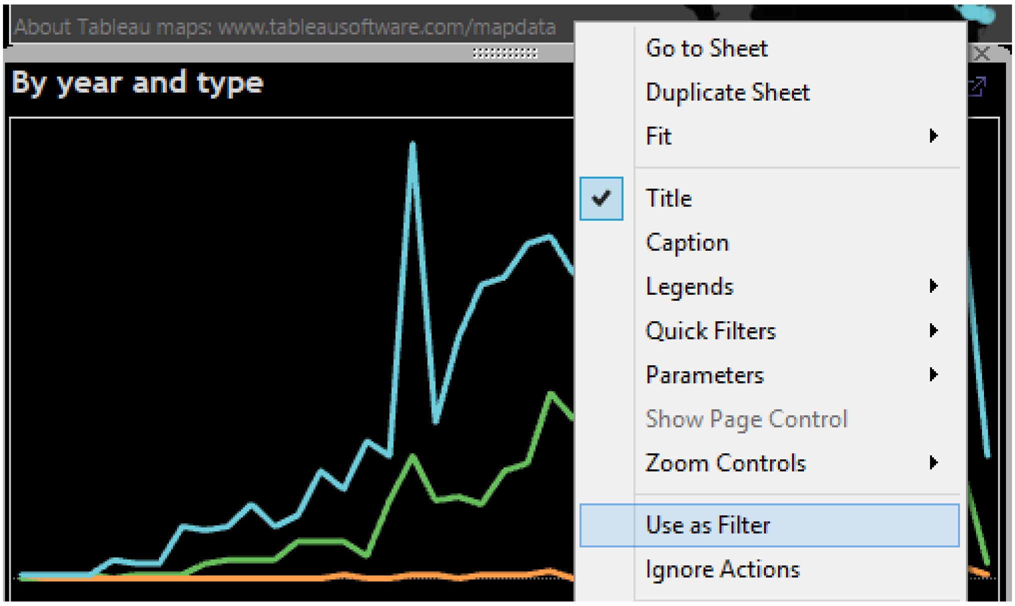


Figure 6.17: Creating filter action

To edit a filter action, we need to access the dashboard's menu option, then select the actions menu to expose the actions dialogue box as shown in Figure 6.18.

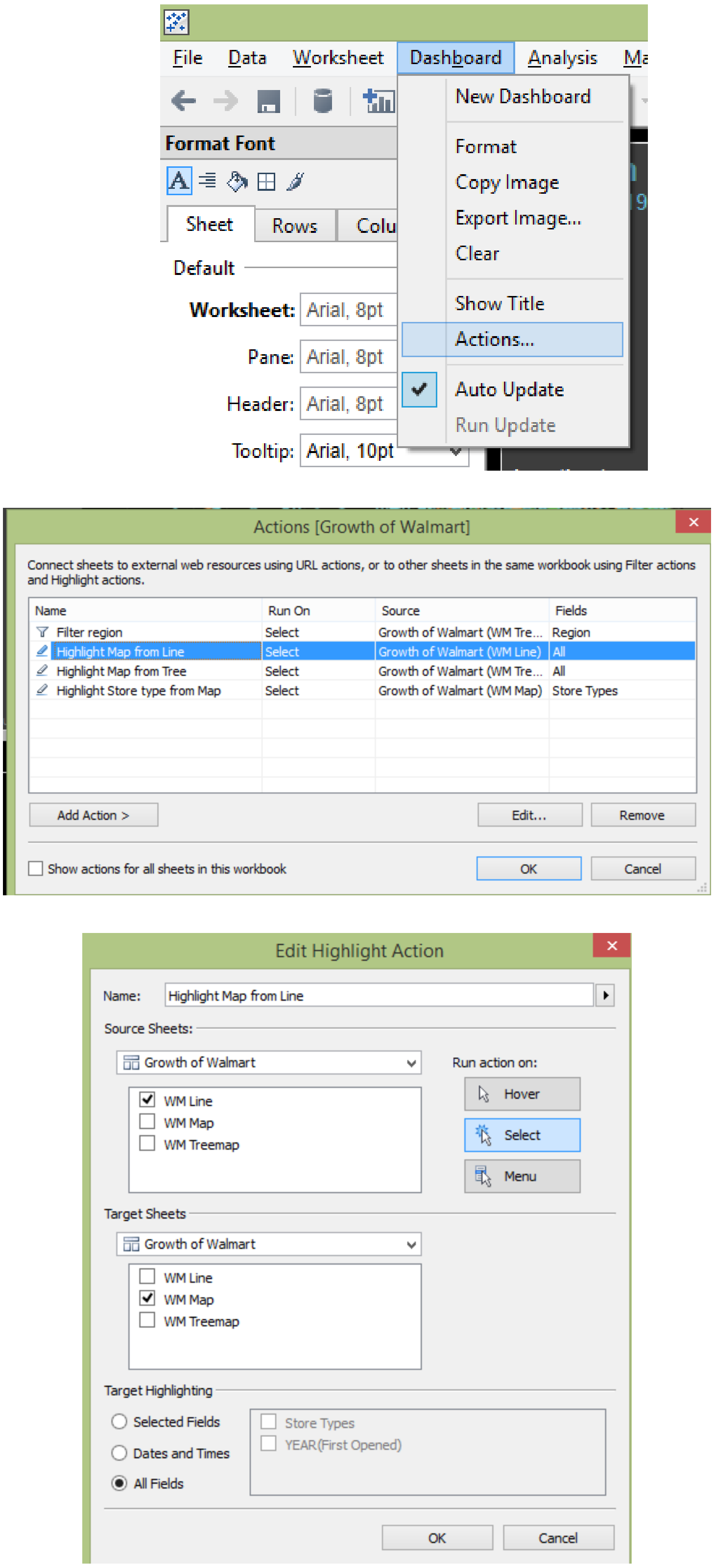


Figure 6.18: Editing filter action

Highlighting helps dashboard viewers to see related information more easily. Users can generate the highlighting function from legends by activating the highlighting tool that appears when we point at the legend, as shown in Figure 6.19.

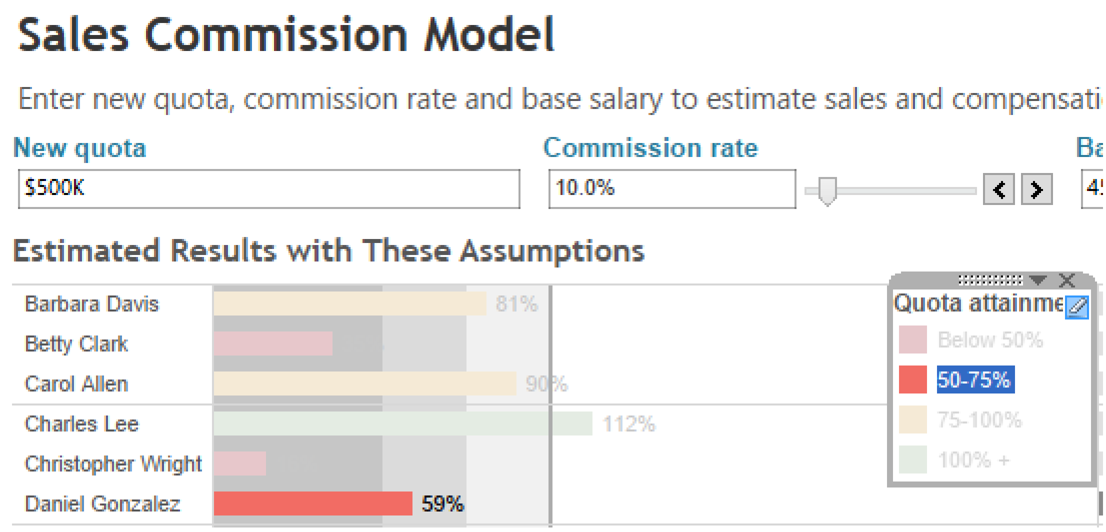


Figure 6.19: Highlighting action

Summary

For sustainability and viability, the automation of scorecard is essential. There are many system functionalities that can be tapped from such automation. Three types of performance dashboards have been introduced: strategic dashboard, tactical dashboard, and operational dashboard. Each type of dashboard serves its users from a different level within the organisation. An overview of performance dashboards with respect to their use and benefits are provided.

This unit also introduced the fundamentals and basic design of a performance dashboard. The fundamental surrounding the design of a performance dashboard has also been discussed. Dashboard design principles are introduced. These features are basic and conceptual. In practice, the dashboard can be an elaborate system that incorporates alert features that can send e-mails, and page or text users whenever the measures are updated.

Whilst an organisation can achieve significant benefits from the initial implementation of a spreadsheet-based scorecard system that is manual. They will eventually feel the need for a more sophisticated system that uses either off-the-shelf software packages or in-house software development. The integration of performance dashboard with the existing enterprise software system can bring many benefits when the entire performance measurement system is unified.

Exercises

ACTIVITY 1

Prepare your answers for the following questions as your instructor may ask you to present your answers in front of the class:

1. Is automating dashboard necessary? How much automation should be carried out?
2. What is (are) the difference(s) among an operational dashboard, tactical dashboard, and strategic dashboard? Explain.
3. Explain the basic dashboard design principles. Why are they important?
4. Which functionalities of the dashboard are the most important and why?
5. How do a dashboard and its automation promote collaboration within an organisation?

ACTIVITY 2

1. Based on the worksheets that you have created in Week 4 and Week 5 class activities, create a dashboard using the Tableau Desktop.
2. Apply the dashboard design principles to ensure your dashboard effectively communicates the information to its viewers.
3. Use filter and highlight actions to help dashboard users and dashboard viewers to see the related information in a dashboard more easily.

Tableau Demos

Dashboard

This demo aims to present how to create Dashboard using Tableau Desktop.

At the end of this demo, you will be able to:
- create a Dashboard

Click here to view the demo.

Action

This demo aims to present how to create Action on Dashboard using Tableau Desktop.

At the end of this demo, you will be able to:
- create Filter Action
- create URL Action

Click here to view the demo.

References

Alexander, M. (2008). *Excel 2007 Dashboards & Reports for Dummies*. USA: Wiley Publishing.

Eckerson, W.W. (2006). *Performance Dashboards: Measuring, Monitoring and Managing Your Business*. New York: John Wiley & Sons.

Eckerson, W.W. (Jul 2006). *Deploying Dashboards and Scorecards*. *TDWI Best Practice Report*. Retrieved on 29 February 2009 from http://download.101com.com/pub/tdwi/Files/RRS_DB_Q206_F_web.pdf.

Few, S. (2007). *Pervasive Hurdles to Effective Dashboard Design*. Visual Business Intelligence Newsletter. Retrieved on 25 February 2009 from http://www.perceptualedge.com/articles/visual_business_intelligence/pervasive_hurdles_to_dd.pdf.

Few, S. (2006, November). *Dashboard Design for Rich and Rapid Monitoring*. *Visual Business Intelligence Newsletter.*

Few, S. (2004, March). *Dashboard Confusion.* Retrieved on 3 March 2009 from http://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf.

Murray, G. D. (2016). *Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software. (2nd Edition)* Indianapolis, IN, USA: John Wiley & Sons Inc.