Author/s:
Hirschberg, J;Lye, J

Title:
Visualising Time: A Compendium of Time Series Plots

Date:
2020-06-01

Citation:
Hirschberg, J. & Lye, J. (2020). Visualising Time: A Compendium of Time Series Plots. The Australian Economic Review, 53 (2), pp.270-291. https://doi.org/10.1111/1467-8462.12374.

Persistent Link:
https://hdl.handle.net/11343/276877

Visualizing Time: A compendium of types of time series plots

Joe Hirschberg and Jenny Lye[1]

Abstract

Graphs are an important for highlighting relationships within a data series or across several series. Modern computer software has provided a flexibility in the construction of graphic displays that would be impossible with the tools available to researchers only a few decades ago. This paper illustrates a variety of different graphical presentations for time ordered or time series data that can now be constructed. These include time series plots, bar charts, range plots, radar charts, scatter plots, heat maps and seasonality plots. For each graph type presented, we discuss the best practice for their construction.

Keywords: time series plots, bar charts, range plots, radar charts, scatter plots, heat maps, seasonality.

1. Introduction

Graphs are an important type of visual display. They can be used to highlight relationships between variables and to identify important values of a series. Graphical methods can be used to show: patterns, trends, seasonality, anomalies, constancy or variation in a series or across series. These methods are not new, in 1786, William Playfair first published "*The Commercial and Political Atlas*" which introduced examples of several graph types that are still in use today (Costigan-Eaves and Macdonald-Ross 1990). Playfair's philosophy was to keep graphs simple, striking and self-sufficient. "*I first drew the Chart in order to clear up my own ideas on the subject, finding it very troublesome to retain a distinct notion of the changes that had taken place. I found it answer[ed] the purpose beyond my expectation…*"(see pg. 320 Costigan-Eaves and Macdonald-Ross 1990).

Today's computer software programs make producing graphs much more straightforward. While one use is in the preliminary analysis of data another is to communicate observations to others. Demir (2015) examined a random sample of articles published in the *American Economic Review* between 1911 and 2010 and concluded that the combined use of graphs is 2.24 counts per article on average.

---

[1] Department of Economics, The University of Melbourne, Victoria 3010, Australia; email <j.hirschberg@unimelb.edu.au>. The authors thank a referee for helpful comments and suggestions.

This is the author manuscript accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1467-8462.12374.

However, a survey of submission guidelines reveals that in most cases there are few guidelines on the construction of graphs. Usually guidelines are restricted to noting that they will appear in black and white in print and must be readable for reduction[2].

Time series graphs are among the most common graphs used in Economics. There is a wide range of graph types that can be used to display time series data. In this paper, we present a range of graph types available and summarize best practice for each graph type.

2. Time Series line Plots

Time series data typically consists of a sequence of data taken at successive equally spaced time intervals. A ***time series line plot*** is a display of the data observations against time along the horizontal axis. These types of plots are useful for examining a single, or a group of time series.

2.1 Single Time Series

In constructing the plot, the *Y*- axis should encompass relevant reference points which are easy to read and indicate the range of the data. Labels on the *Y*-axis should be centered to the left of the appropriate tick mark and as close as possible to it. Similarly, labels on the *X*-axis should be centered beneath the appropriate tick mark and as close as possible to it. The weight of the line should not be so thin it fades into the background but not so thick it obscures data points. Inner gridlines should only be used if you want the reader to extract specific values. If they are to be used, they should be visible but light, for example gray and with a delicate line. An alternative to gridlines is to include a second *Y*-axis. Some packages include backgrounds by default. However, they can make graphs difficult to read. Graphics on a white background are the most legible.

A time series plot can help highlight characteristics of the data. For example, it may reveal any oddities such as outliers and discontinuities or features such as trend and seasonality or it may indicate that a variable requires a transformation prior to the main analysis.

Stationary time series

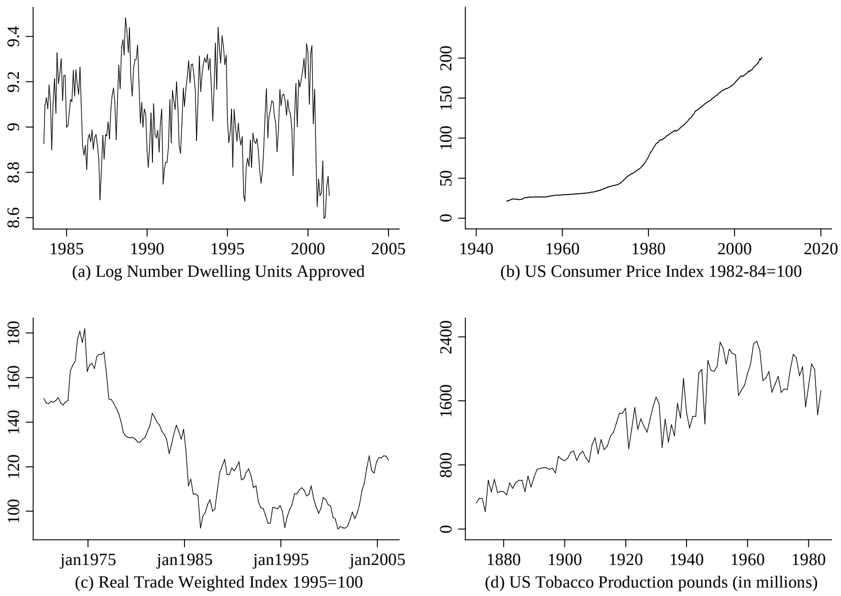
Stationary time series have: no systematic change in mean (no trend), no systematic change in variance and when strictly periodic variations have been removed. This implies that the properties of one section of the data are much like those of any other section. A time plot of the series is useful to detect if the mean and the variance remain approximately stable over time.

Several time series plots are illustrated in Figure 1. Figure 1a plots monthly data on the log of number of building approvals for private houses. The time series appears to fluctuate around a constant value, it also seems to have no discernible

[2] Note that the American Economic Association Journals publish all papers with colour graphics. Also, the online versions of many journals are provided in colour.

pattern or trend and is consistent with the behaviour of a stationary variable. Figures 1b, 1c and 1d illustrate variables that display characteristics of nonstationary behaviour. Quarterly data on the US CPI is plotted in Figure 1b and displays upward trending behaviour. Figure 1c plots quarterly data on the real weighted trade index of the Australian dollar which displays meandering behaviour. The yearly US tobacco production plotted in Figure 1d exhibits an overall trend. The fluctuations increase as the level of the series increases which indicates that the variance of the series is time dependent[3].

Figure 1: Stationarity Properties of Time Series



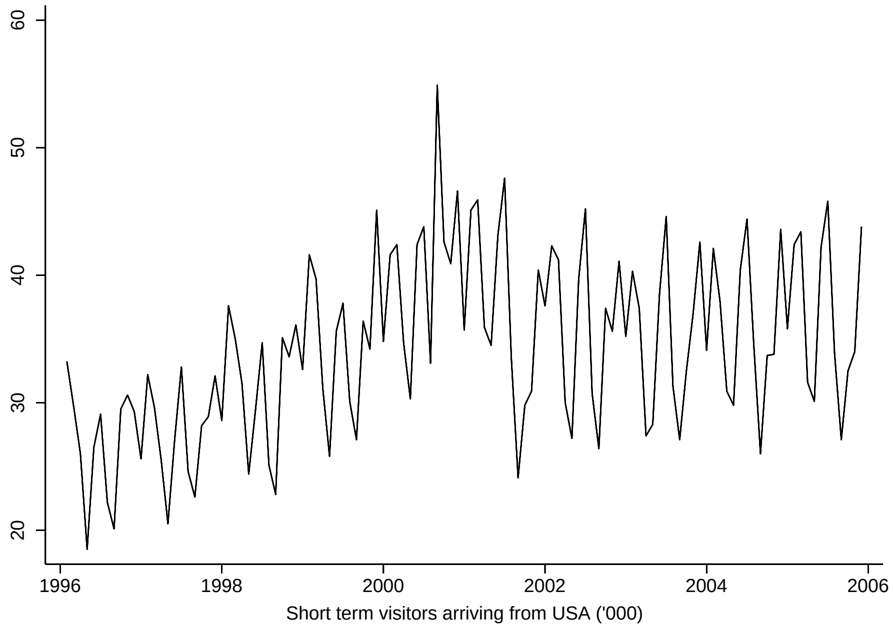
Seasonal Time Series

Seasonality is cyclical behaviour that occurs on a regular calendar basis and maybe due to factors such as weather which may influence tourism and construction or specific cultural or religious events such as Christmas and New Year. Figure 2 illustrates an example of a series that contains a regular seasonal pattern. Monthly data on the short-term visitor arrivals to Australia from the US is plotted. While a regular seasonal pattern can be observed in the plot, we can also note that visitor numbers from the US were increasing up until the 2000 Sydney Olympics and then appeared to have plateaued afterwards[4].

[3] Sources of data include: Figure 1a - ABS 8731.0; Figure 1b - Department of Labor: Bureau of Labor Statistics; Figure 1c - Reserve Bank of Australia FRERTWI and Figure - 1d Wei (2006).

[4] Data source: ABS 3401.0.

Figure 2: Seasonal Time Series

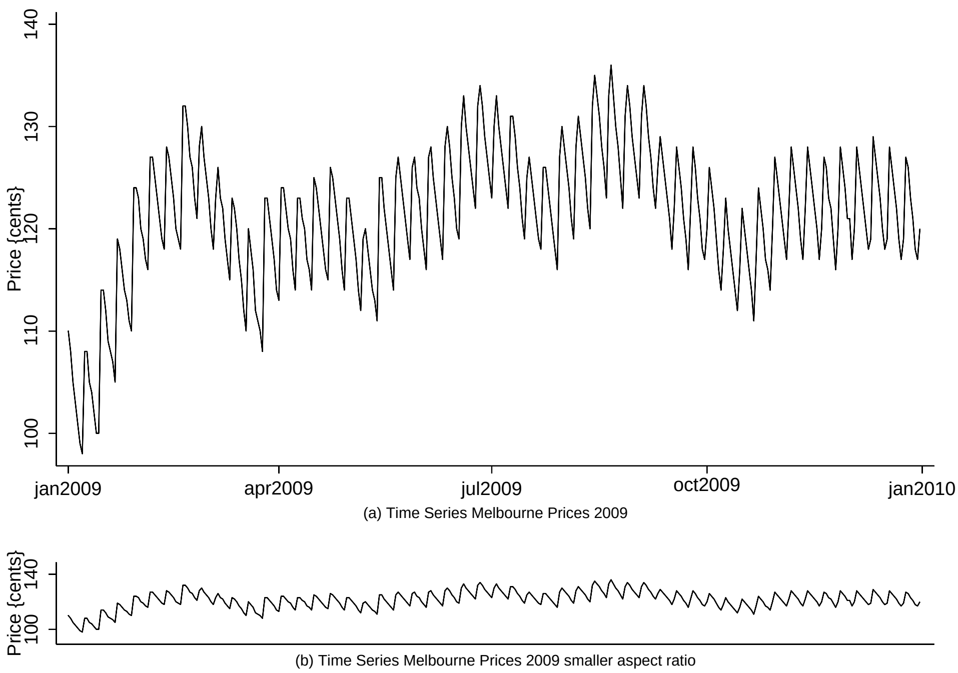


The Aspect Ratio

The aspect ratio is the ratio of a graph's height to width. For cyclical data, changing the aspect ratio can help identify the changes in the peaks and troughs over time. Figure 3a plots daily retail petrol prices in Melbourne in 2009 that presents obvious cyclical behaviour. Figure 3b is the same plot but with a smaller aspect ratio. This figure highlights how the cycles change with a faster increase than their decline[5].

[5] Australian Institute of Petroleum Melbourne Daily retail price cents per litre inclusive of GST (BP Australia, Caltrex Australia Petroleum and Mobil Oil). Numbers digitized of graphs provided. *http://www.aip.com.au/*

Figure 3: Illustrating the Aspect Ratio

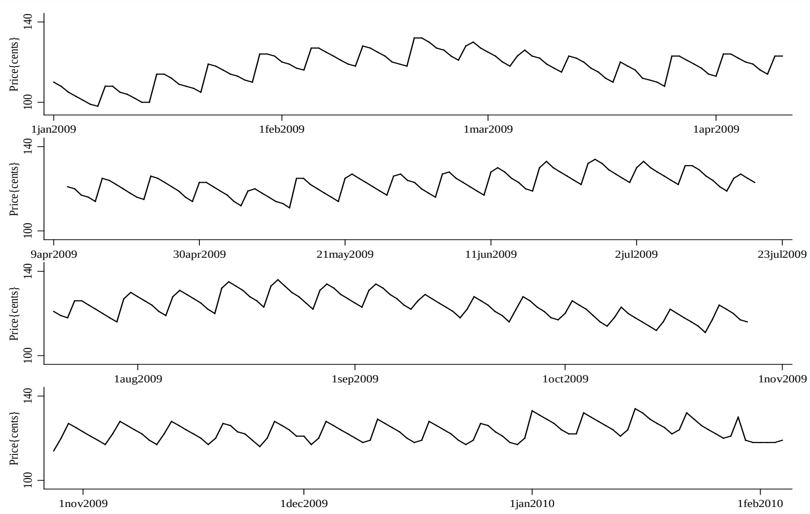


(a) Time Series Melbourne Prices 2009

(b) Time Series Melbourne Prices 2009 smaller aspect ratio

A long time series will tend to have a small optimum aspect ratio. However, this may lead to a graph that is difficult to read. One alternative is to: break up the horizontal axis into several intervals, graph the subset of data and then stack the graphs. Figure 4 plots the retail petrol data using this technique. Again we can clearly see the pattern in the cycles although now the plots are more spread-out[6].

Figure 4: Slice Plot of Melbourne Petrol Prices

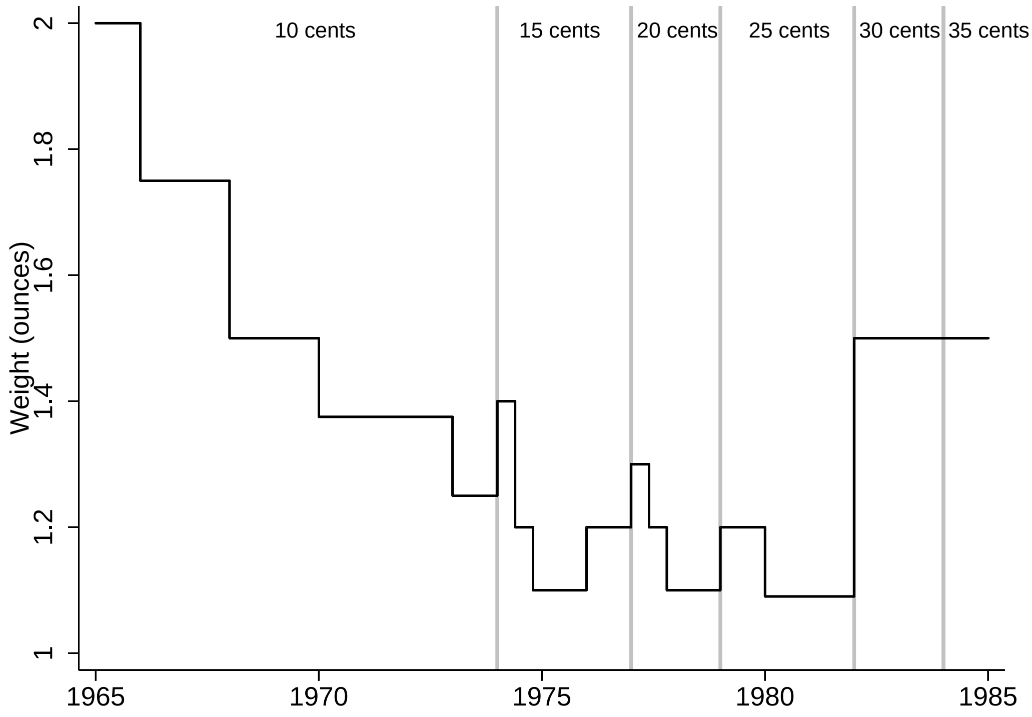


[6] The source is the same as Footnote 5.

Reference Lines

Reference lines are a useful addition to the plot to highlight an important value. They are usually shaded a lighter colour so as not to obscure the data line. Figure 5 plots the weight of a Hershey bar with reference lines indicating price changes. The reference line helps to highlight a decline in the size of bar with a fixed price followed by a sudden increase of price and size (see also Cleveland 1985 pg. 188). The step nature of the plot is caused by prices that remain constant over varying intervals.

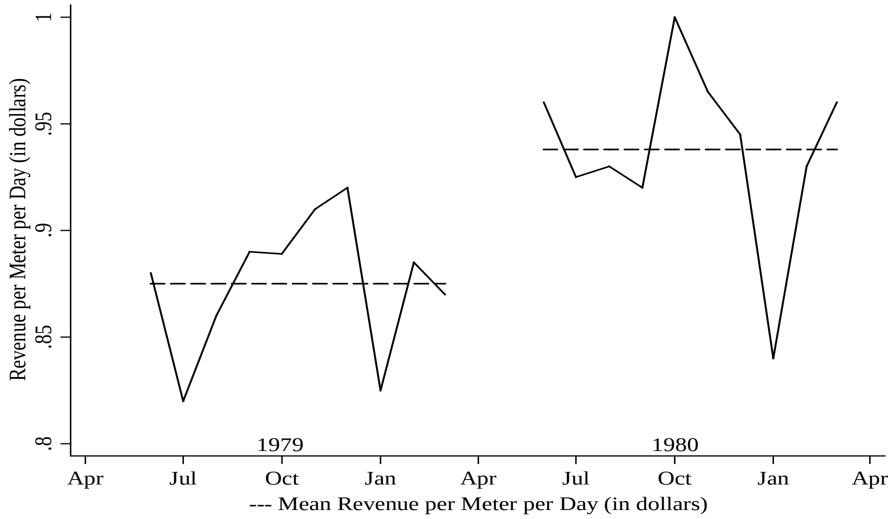
Figure 5: Weight and Price Changes



Finding Data Anomalies

Time series plots can be used to highlight anomalies in the data. In the late 1970s the City of New York owned and operated around 70,000 parking meters from which an average of $50,000 was collected daily. In May 1978 Brink's Inc was awarded a contract to collect the coins. On April 9, 1980 five Brink's collectors were arrested for theft of coins from the parking meter collections. On April 9th a replacement firm (CDC) took over the collection process and measures were taken to ensure that there was no theft by employees. Figure 6 plots daily revenue per meter (see Finkelstein and Levin 2001, pg 7). Two consistent ten-month periods are plotted – one covering the Brink contract and the other the CDC contract. Also included in the Figure is the mean revenue per day collected given by the dashed line. Figure 6 clearly highlights the difference in the amount collected between the two firms and provides evidence of suspicious behaviour in the collection process in the first 10 months.

Figure 6: Parking Meter Revenue

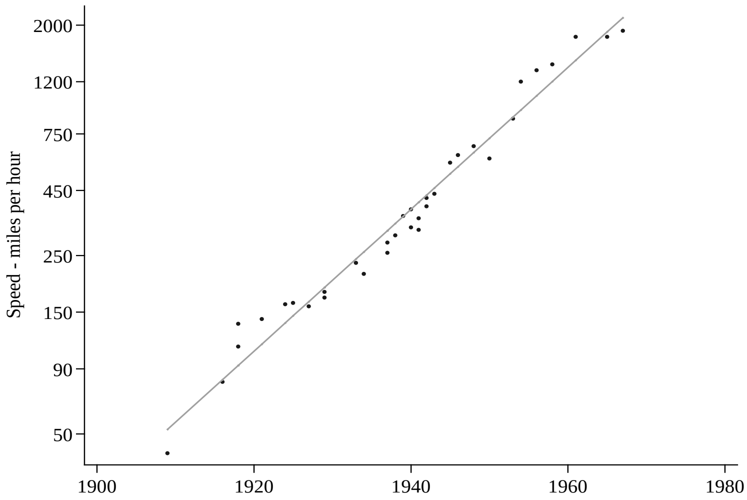


Semi-Logarithmic Plot

A semi-logarithmic plot, also known as a ratio scale graph, is used to study the rate of change of a certain variable over two different points in time. The interpretation depends on the slope of the curve, whether ascending or descending. The slope of the lines in these plots can be interpreted as a percentage change or rate of change. This means that a curve increasing at a constant rate is a straight ascending line and decreasing at a constant rate is a straight descending line. An ascending convex curve indicates an increase at a decreasing rate whereas a curve increasing at an increasing rate is concave. A curve that is decreasing at a decreasing rate moves downward in a concave fashion whereas it is decreasing at an increasing rate if it is convex downward. A horizontal curve is neither increasing nor decreasing.

Figure 7 plots the top speed of US combat aircraft from 1909 until 1970 (see Martino 1971) using a semi-logarithmic chart including a superimposed time trend line. Over this time period there were several innovations to these types of aircraft including the closed cockpit, monoplane, all metal airframe and the jet engine. However, despite these varied types of innovations the rate of progress of the speed trend appears to be quite steady.

Figure 7: Top Speed US Combat Aircraft



2.2 Multiple Time Series

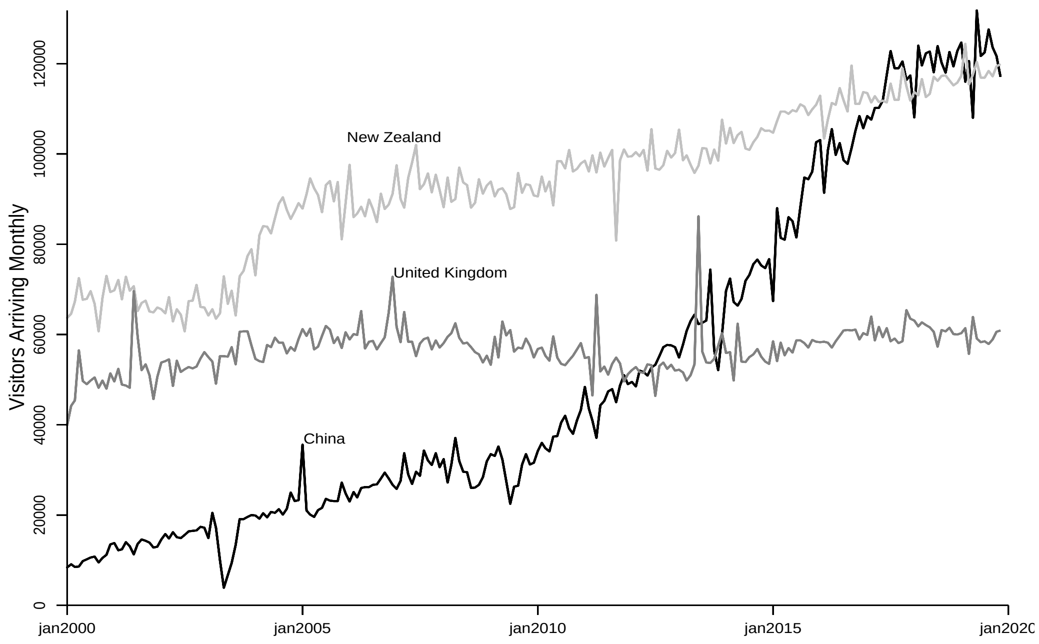
Groups of variables are often plotted on the same graph. It is highly recommended to plot no more than four variables on a single graph. Such plots may reveal possible relationships between the different series or be used to highlight different features of the data. In such a plot the lines can be differentiated by varying weights and shades rather than adding markers which can be a distraction and interrupt the visual flow. A legend should be placed as to not interfere with the interpretation of the chart. The usual practice is to position the key in the top right corner unless space is available only in another part of the display. Alternatively, concise labelling of lines makes it easier to compare patterns. In an experiment conducted by Milroy and Poulten (1978) they found that readers could use direct labels more quickly than keys without loss of accuracy.

Multiple Line Plots

Figure 8 plots the number of visitors arriving monthly in Australia from New Zealand, the UK and China over the period 2000 to 2019. While over this period the number of visitors arriving from the UK appears to be relatively stable, increases can be seen for both visitors from New Zealand and China. With the increase from China to be at a much faster rate since 2010[7].

[7] Data source: ABS 3401.0.

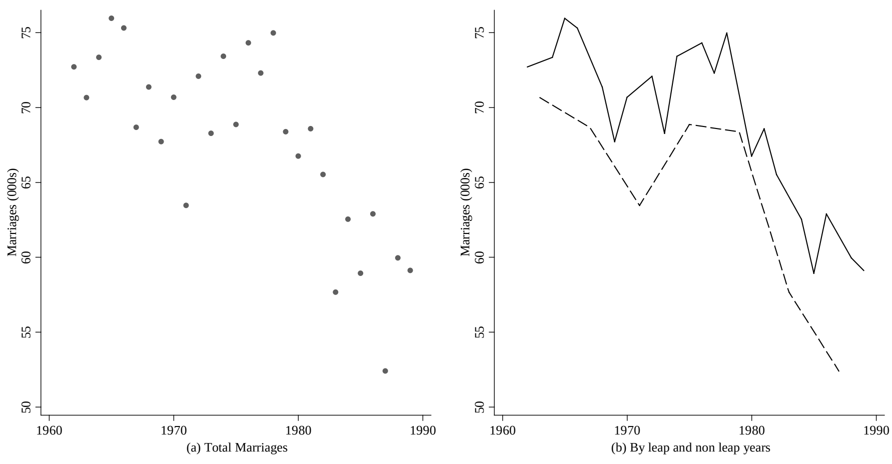
Figure 8: Monthly Visitors to Australia



Multiple Plots of Same Series

A symbol plot of a time series may be used to find the long-term trend. Connecting the points can reveal cycle patterns in the data. Figure 9a is a symbol plot of the number of marriages per year for Greece over a 22-year period 1968-1990 (see Pole et al. 1994). There is a downward trend in the annual data but there also appears to be a cyclical pattern which is unusual in annual data. In Figure 9b we separately connect non-leap years (-) and leap years (--) and it is now clearly seen that it is the Greek tradition of not being married during a leap year explains the cyclical nature of the data.

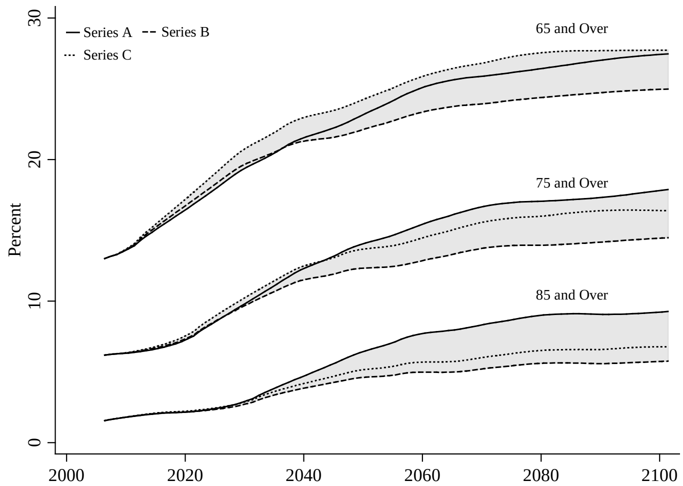
Figure 9: Marriages in Greece



Multiple Projection Comparisons

When plotting a time series projection one can use distinctive patterns to identify the separate projections with shadings to emphasize variations between highest and lowest projections. Figure 10 plots three alternate Australian population projections for those over 65 years, those over 75 years and those over 85 years from 2006 to 2101. The projections correspond to three scenarios, A (high levels of fertility, mortality, net overseas migration and net interstate migration); B (medium levels) and C (low levels). Each series is indicated by a different pattern defined in the legend. We see that the percent of older population is forecasted to increase over time in all three projections, but they appear to plateau to different levels. The shading highlights that the further the projections extend into the future the wider the variations[8].

Figure 10: Population Projection for Older Ages



Multiple Area Plots

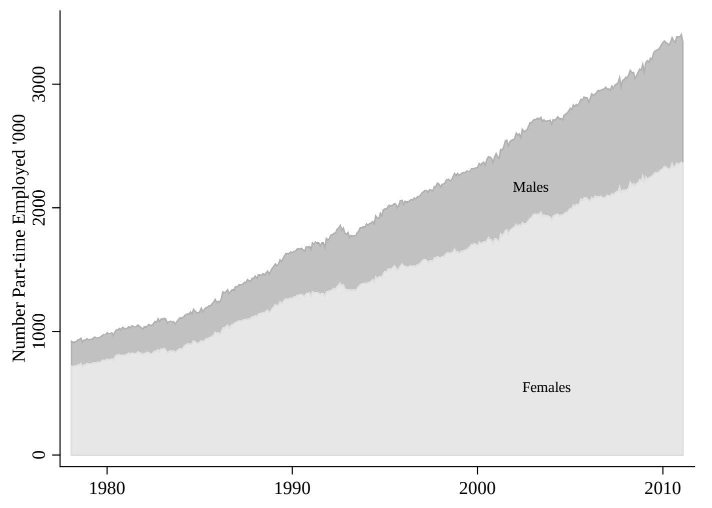
Time series multiple area plots are like time series line plots except the area under the line is shaded. This chart should include a zero baseline. With two or more series, the series need to add up to a whole and the basis for shading is to highlight the distance between the plotted lines rather than the distance of the lines from the zero baseline. The various areas should be distinctively shaded. Figure 11 plots two time series[9]. The first represents the number of part-time female employees and the second line corresponds to the total number of part-time employees. Both sets of data are seasonally adjusted. The difference between the two lines corresponds to the number of part-time male employees. We see that the number of part-time employees has been increasing over time. The shading indicates that while this is mainly due to an

[8] Data source: ABS 3222.0.

[9] Data Source: ABS 6202.0.

increasing number of part-time female employees there has also been an increasing number of male part-time employees.

Figure 11: Number of Part-Time Employees



Sparklines

Sparklines are "*intense, simple*, *word-sized graphics*" that can be used to add context to data in a sentence or embedded in a cell in a Table (see Tufte 2006). They are a useful way of tracking changes over time or can be used to allow comparison across different groups and are often added to tabular displays.

Figure 12 is an illustration which includes time series plots of seasonally adjusted quarterly unemployment rates for Australia covering the two recessionary periods 1981-83 and 1989-93[10]. In this example, each time series plot was constructed individually where the scale of the *Y* axis on each plot was the same. Figure 12 is constructed by inserting each plot into the appropriate cell of a Table in a Word document. The sparkline is kept simple by removing the clutter of the *Y* and *X* axes labels and keeping the trend of relative values. A comparison across the different time series plots clearly illustrates that for each of the States the two recessionary periods are associated with higher levels of unemployment although there is a great deal of diversity across the States.

[10] Data source: ABS 6203.0. Note that monthly figures are averaged to obtain quarterly figures.

Figure 12: State Unemployment Rates

State	***1978q2***		***1998q4***	***Min***	***Max***
NSW	6.3		6.8	4.9 1981*q*2	11.1 1983*q*3
VIC	5.6		7.4	4.4 1989*q*4	12.0 1993*q*3
QLD	7.1		8.2	4.6 2005*q*1	10.6 1993*q*3
SA	6.7		9.1	4.8 2005*q*3	11.8 1992*q*2
WA	6.3		6.9	4.1 2005*q*4	11 1992*q*1
TAS	6.4		10.1	5.5 1979*q*2	12.5 1993*q*3

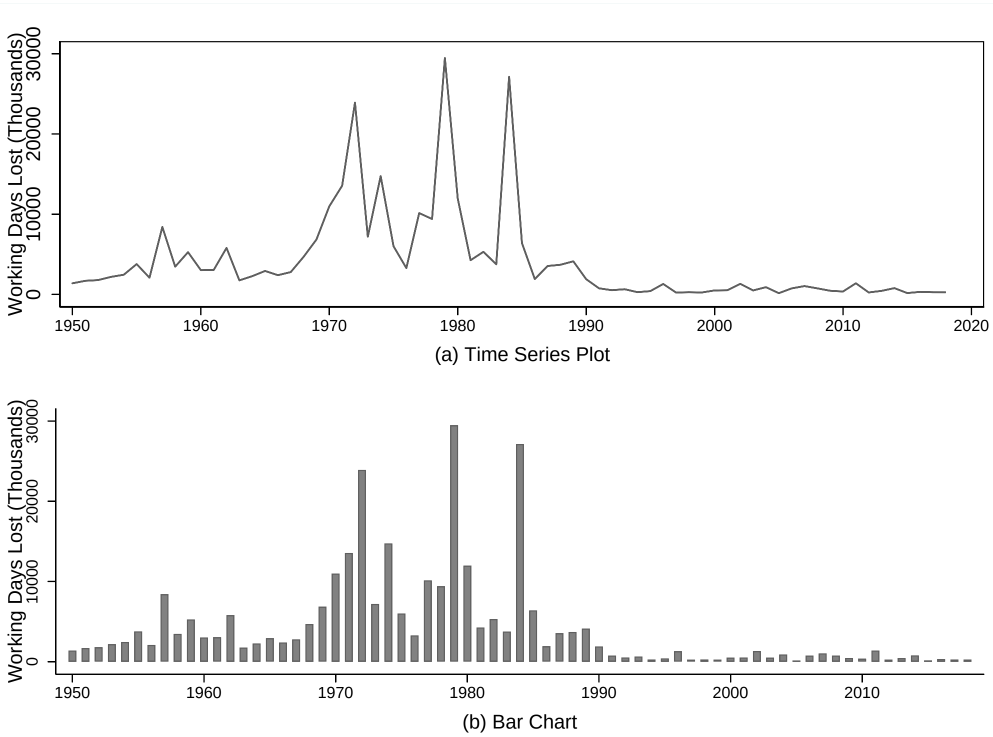
3. Bar Charts

For time series data vertical bars are used with the length of each bar representing the quantity to compare and the higher or longer the bar, the greater the value of the variable in the time series. Bar charts allow the viewer to see the individual values. The bars should be of uniform width and evenly spaced. The vertical axis should always begin with 0 and cover the range of the data.

Simple Bar Chart

Figure 13a is a time series plot of working days lost per year due to industrial disputes in the UK over the period 1950-2019[11]. Figure 13a highlights the fluctuations over the years although in the last two decades the number of working days lost per year due to industrial disputes has decreased dramatically. To emphasize the size of the fluctuations an alternative approach is to plot the data using a column chart as in Figure 13b.

Figure 13: Working Days lost due to Industrial Disputes UK



Scaled Bar Chart

Bar charts can also be used to emphasize positive and negative numbers, increases and decreases and gains and losses. In this case the bars can extend either above or below the zero line. Figure 14 plots the percentage changes in US seasonally adjusted real GDP based on quarter to quarter changes[12]. Shading between the reference lines marks the beginning and end of recessions as determined by the Business Cycle dating committee of the NBER. The financial press often states the definition of a recession as two consecutive quarters of decline in real GDP. Most but

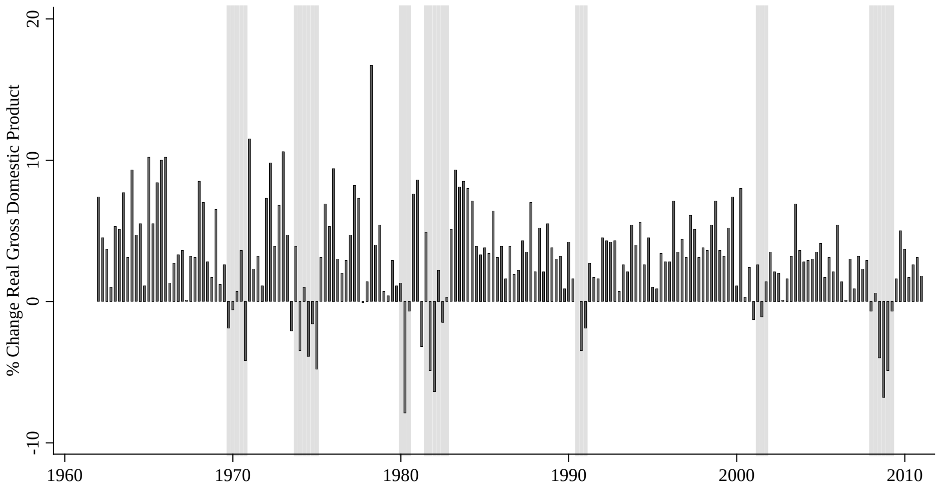
---

[11] Labour disputes in the UK: Office for National Statistics.

[12] Table 1.1.1 Percent change from preceding period in real GDP. National Income and Product Accounts Table. Bureau of Economic Analysis. Beginning and end recessions determined by the Business Cycle Committee of the NBER.

not all of the recessions identified by the NBER consist of two or more quarters of declining real GDP.

Figure 14: Percentage Changes in US Seasonally Adjusted Real GDP

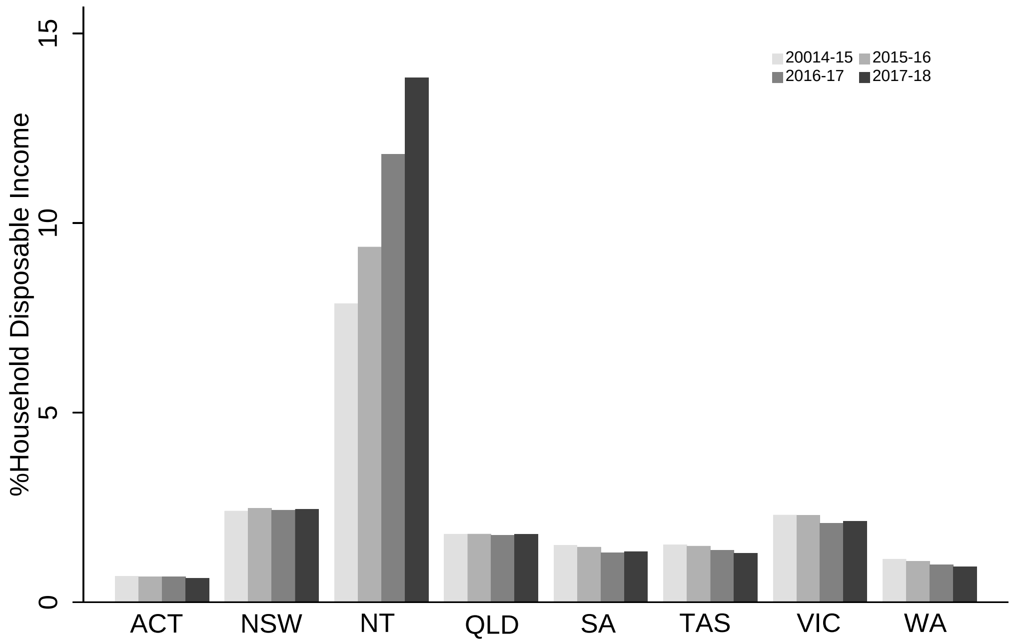


Multiple Bar Charts

Multiple bars representing different series can be graphed together and this is another effective method of comparing sets of data. However, too many data series will create clutter. Shading of the multiple bars should move from lightest to darkest.

Household disposable income is a key determinant of how much households can spend on consumer goods and services. Figure 15 plots the proportion of household disposable income spent on all gambling forms for each Australian state and Australia as a whole for the fiscal years 2014/15 to 2017/18 (see AGS 2019). Each bar plotted represents a different year. For most states and for all of Australia, there has either been a decline or at least no increase in the proportion of household disposable income spent on gambling. The exception is the Northern Territory which has the largest percentage as well as an increasing percentage over time. This anomaly has been caused by the number of online providers that have been attracted to the Northern Territory by lower taxes (see Barnes et al. 2017).

Figure 15: Total Gambling Expenditure as a Percentage of Household Disposable Income

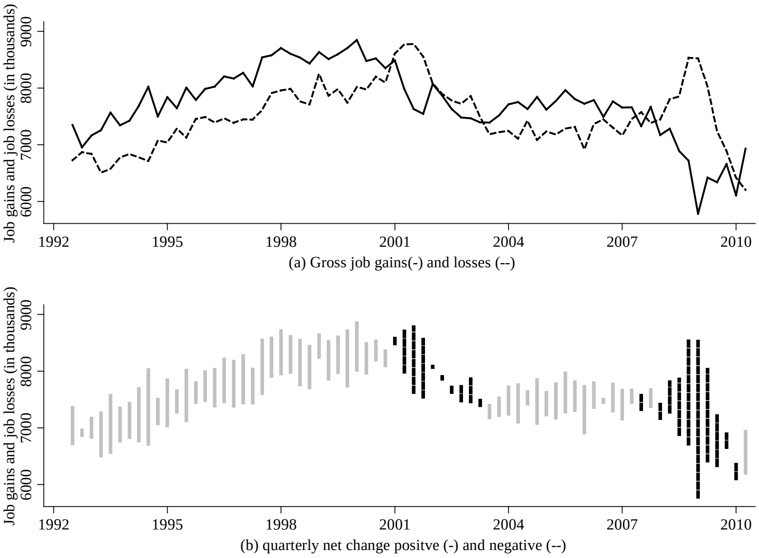


4. Range Plots

Sometimes, it is the difference between two series that is of interest such as the difference between high and low daily stock prices or the range. Range plots can be used to illustrate such a difference. The length of each bar indicates the fluctuations from the low and high values. Shading of the bars can distinguish between positive and negative fluctuations. Figure 16a is a time series plot of US private sector gross job gains and losses[13]. The range plot in Figure 16b can be used to highlight the net change - the difference between gross gains and losses. In Figure 16b, the length of the bar represents the net change and bars are shaded differently to highlight whether the net change is positive or negative (see also Carson 2009).

[13] US seasonally adjusted private sector gross job gains and gross job losses US Bureau of Labor Statistics.

Figure 16: US Private Sector seasonally adjusted gross job gains and losses



5. Radar Chart

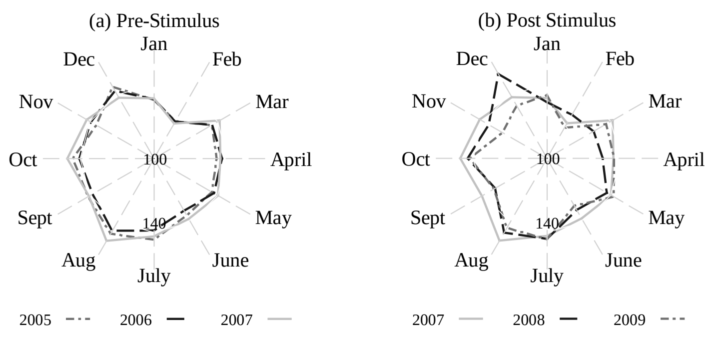
A radar chart is a plot of circular time series. These are values that are related by a repeating cycle. These plots have a separate axis for each category and the axes extend outward from the centre of the chart. For example, monthly data are circular where the different months are seen as separate categories with 12 axes but they are related by their order. The value of each data point is plotted on the corresponding axis. The chart allows easy comparison among data points and is useful for showing outliers and commonality.

As a response to the global financial crisis of 2008/09 the Australian Government announced two economic stimulus packages that included payments to households. The first, December 2008 gave cash bonuses predominantly to low income families and pensioners. The second, included cash bonuses to Australian taxpayers and were paid over April and May 2009. This package also included payments to eligible families. There was some suggestion in the media that these packages were associated with a splurge on spending on the pokies. In Figure 17 we plot monthly real net expenditures on gaming machines in Victoria (1989/90=100)[14]. The first radar plots the 3 years prior to the stimulus packages to show how consistent the spending on the pokie machines had been. The second radar plot includes 2007 and the two years associated with the stimulus payments 2008 and 2009. There is a clear increase in December 2008 spending which coincides with the payments for the first stimulus package. Similarly, spending in March and April 2009 was similar to the 2007 figures whereas all the other months had generally lower expenditures.

---

[14] Monthly gambling expenditures Victoria *www.vcglr.vic.gov.au*.

Figure 17: Gambling Expenditure on Pokie Machines in Victoria



6. Scatter Plots

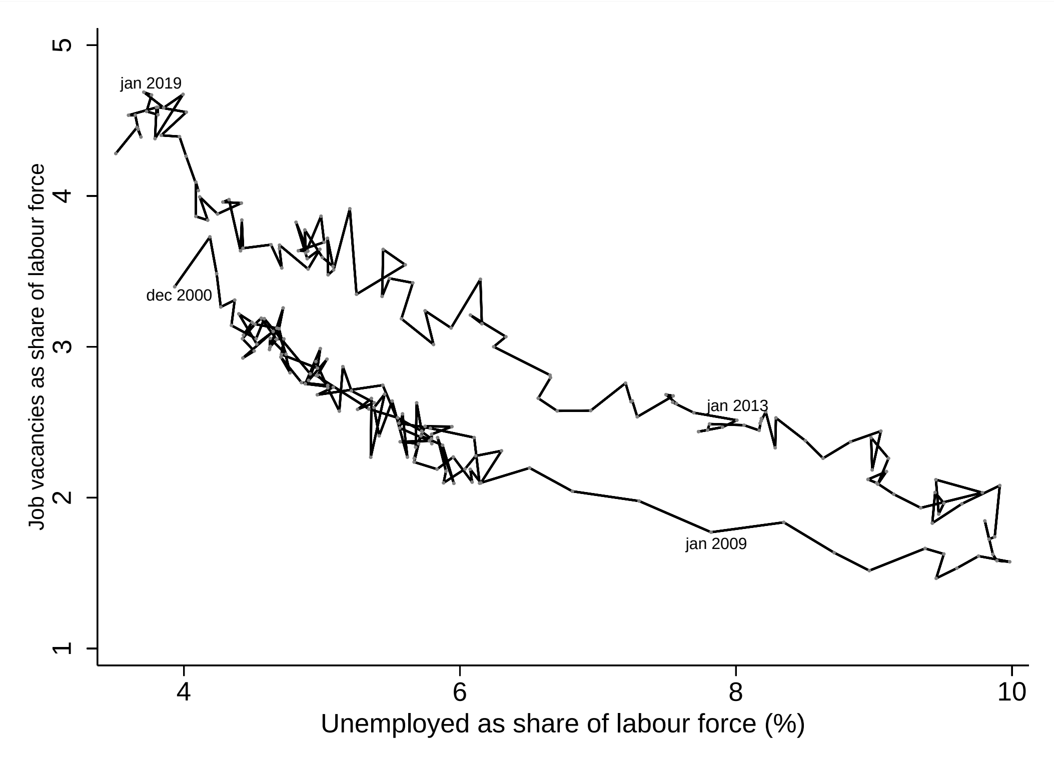
Scatter plots show the relationship between two variables. If markers are to be used to identify points in the scatter plot, they should not be too large so that they are plotted on top of each other and hence obscuring any relationships between the variables. This can be avoided by making the markers smaller or using hollow glyphs.

Time Connected Scatter Plot

To indicate a sequence of events and to also highlight distinct patterns in the data the points on a scatter plot can be connected and in this way we add a third dimension to the scatter plot. A Beveridge curve shows the relationship between unemployment and job vacancies. Figure 18 plots this relationship for the US[15]. We can see that the job vacancy rate fell and unemployment increased during the Global Financial crisis. More recently, the Beveridge curve has shifted right and upwards indicating that job vacancies have increased coinciding with a fall in the unemployment rate.

[15] Data from St Louis FRED: Total Unfilled Job Vacancies for the United States and Total Unemployed, Plus All Persons Marginally Attached to the Labor Force, Plus Total Employed Part Time for Economic Reasons, as a Percent of the Civilian Labor Force Plus All Persons Marginally Attached to the Labor Force.

Figure 18: The Beveridge Curve for US



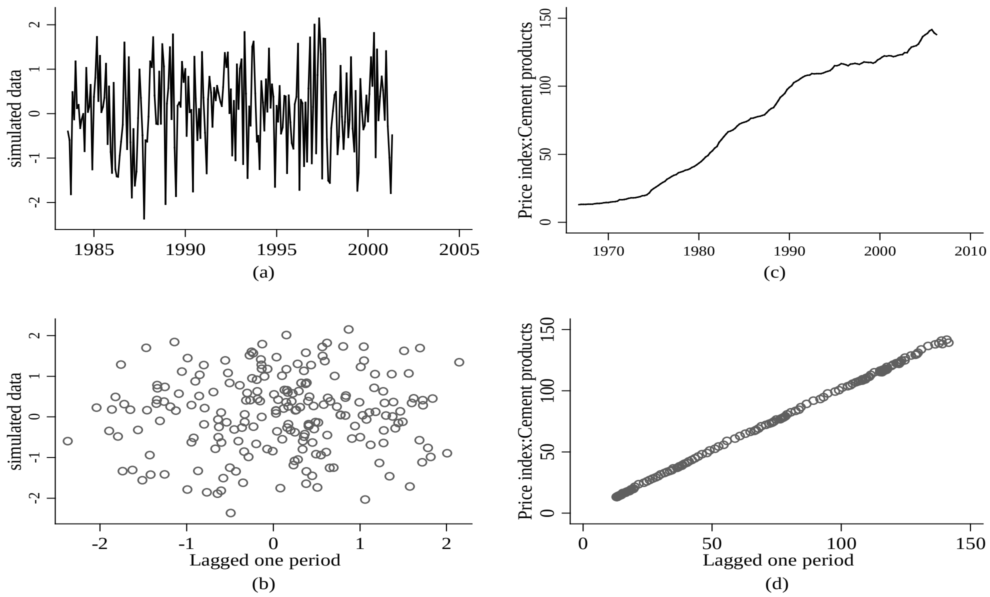
Lag Plots

Correlation measures the degree of dependence or association between two variables. Autocorrelation means that the value of a series in one time period is related to the value of itself in previous periods. One way to detect autocorrelation in a series is by the construction of a scatter plot of a series versus lagged values of the series. These plots are referred to as lag plots.

Figure 19a plots simulated random data. The lag plot of the series against its first lag is plotted in Figure 19b and has the appearance of a shotgun pattern which indicates completely random data. Figure 19c is a time series plot of quarterly data of a price index of cement products used in house building in Australia and the lag plot of this series against its first lag is presented in Figure 19d[16]. From Figure 19d we see a tight clustering of points along the diagonal. Such a pattern indicates strong positive autocorrelation. The series is highly non-random and there is strong association between an observation and a succeeding observation.

[16] Data source: ABS 6427.0.

Figure 19: Time Series data and Lag Plots

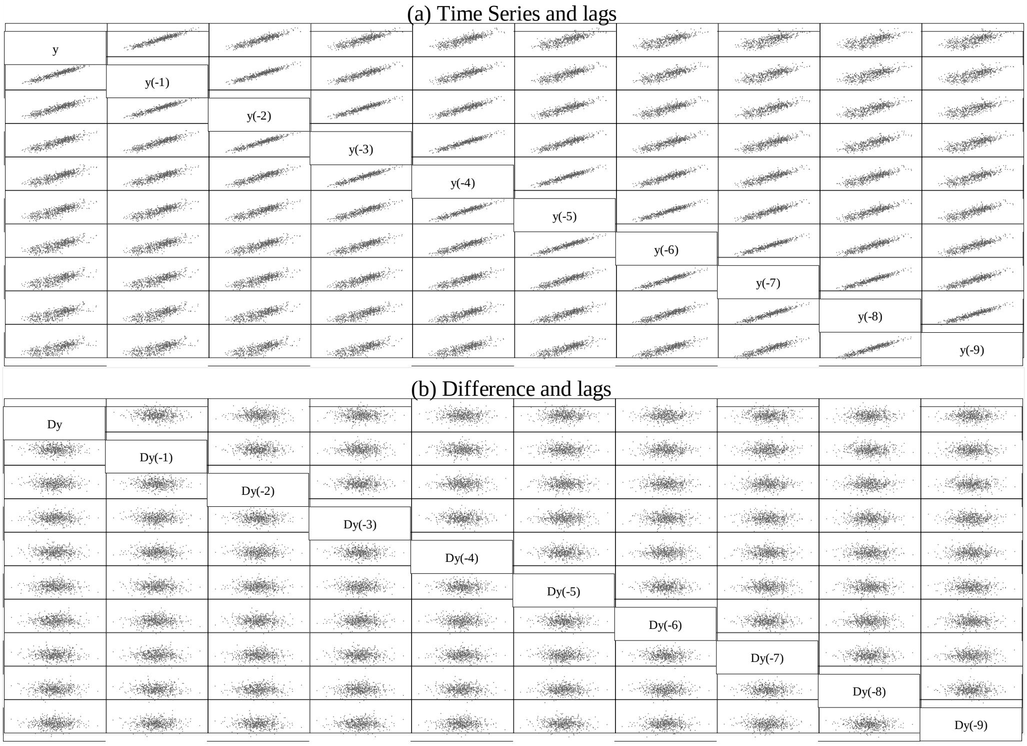


Matrix Lag Plots

A matrix scatterplot consists of an array of scatter plots such that any adjacent pair of plots have an axis in common**.** It presents the data with little embellishment or extra graphical features in order to emphasize patterns of data across the cells of the matrix**.** While any number of variables can be included in such a graph it should be remembered that it is often easier to interpret a graph if it can fit onto a single page**.**

Figure 20a shows a matrix lag scatterplot of a random walk series and the first nine lags of the series**.** The figure highlights a very clear pattern that is consistent with a random walk series. Close to the diagonal the correlations are very high but these decrease uniformly as we move away from the diagonal**.** Figure 20b shows the corresponding scatterplot matrix of the first difference of the random walk and the first nine lags of this differenced series**.** This matrix scatterplot highlights that the differenced series is white noise with each scatterplot off the diagonal displaying a shotgun pattern.

Figure 20: Matrix Scatter Plot of Time Series, First Difference and Lag Plots



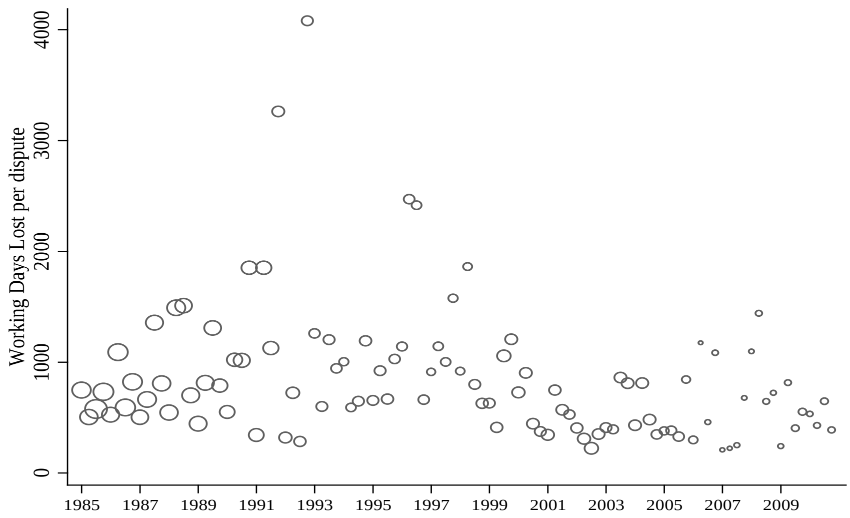
Bubble Plots

Another method of adding an extra dimension to a scatter plot is to change the plotting symbol. This type of scatter plot is known as a bubble plot where each point is represented by a circle or “bubble” that varies in size based on the values in a third data series. Typically, larger values of the third series are shown with larger bubbles. Open circles or bubbles are used rather than a filled, plotting symbol because usually there will be some degree of over plotting in the display. Figure 21 plots quarterly data on working days lost in Australia per dispute from

1985 until 2010[17]. Each point in the plot is represented by bubble and the size of the bubble is determined by the number of employees involved per dispute. Generally, apart from a few outliers, the number of working days lost per dispute has been relatively constant. However, what the bubbles highlight is that the number of employees per dispute has fallen.

[17] RBA Australian Economic Statistics 1950-1995 and ABS Source 6321-01 1958-2010. Annual data obtained from summing quarterly data.

Figure 21: Working days lost in Australia per dispute



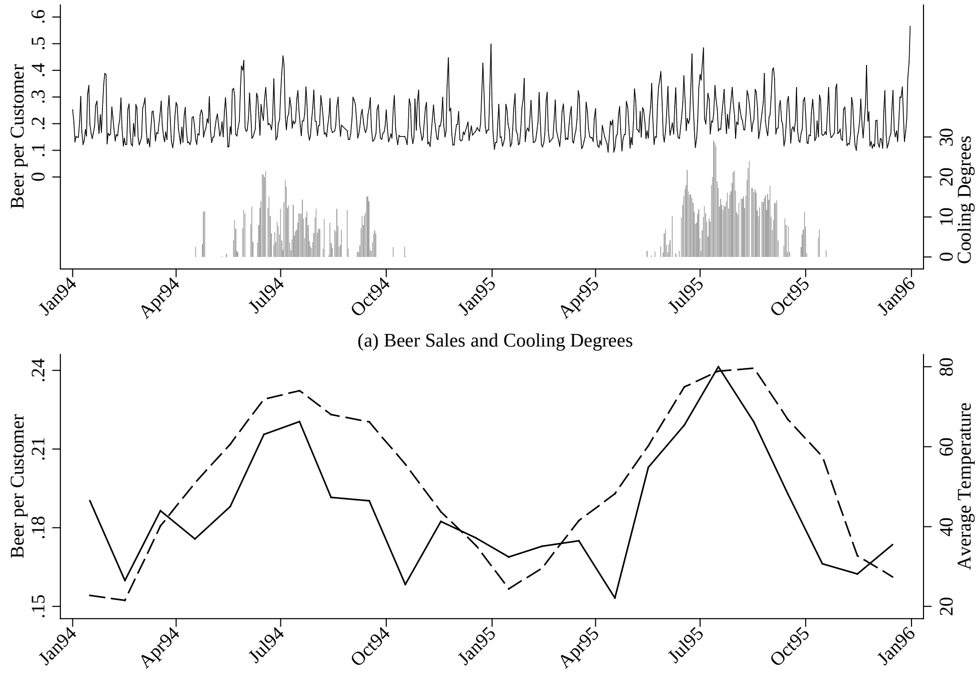
7. Multiple Y-Axis Scales

Multiple series are sometimes plotted on the same plot where the series share a common *X*-axis but have different *Y*-scales. These types of plots should normally be limited to only two different *Y*-scales. This involves using two different vertical axes on each side of the plot which are defined in different units. They are useful for comparing curves that vary widely in magnitude or in different measure. The two different *Y*-scales can be drawn to either overlap or not overlap each other. The two scales should not overlap in the case when the different plots intersect each other, and it becomes difficult to distinguish them. If the scales are over-lapping, then the division of each of the scales should be spread in the same manner.

The Dominick's Fine Foods data includes daily sales by store for a chain of food stores in the Chicago metropolitan area by departments[18]. In this example we examine two years of daily data from one of the stores to determine if there is a relationship between the per customer sales of the beer department and the weather. The weather measures are the average temperature and the number of cooling degrees for each day. Cooling degrees are defined as the number of degrees greater than 65 degrees Fahrenheit registered for the average daily temperature and are used as an indication of the need for air-conditioning. In Figure 22a the cooling degree days indicate when the higher beer consumption occurs however there is a lot of noise in the data that masks the relationship. Most of this is in the form of day to day demand fluctuations. In Figure 22b we use an alternative plot where we plot the median value of beer sales by month and a smoother indicator of weather variability defined by the average temperature. This Figure clearly indicates the relationship between beer consumption and temperature.

[18] See https://www.chicagobooth.edu/research/kilts/datasets/dominicks.

Figure 22: Sales of Beer and Weather



(a) Beer Sales and Cooling Degrees

(b) Beer Sales (-) and Average Temperature (--)

8. Seasonality

Seasonality is a pattern that repeats itself each year. Over time this pattern may drift or change in amplitude. While a time series plot will show periodic behaviour it is difficult to determine the nature of the seasonality from this plot. Two useful plots to highlight the nature of the seasonality are the cycle plot and the vector of periodicity plot. A cycle plot shows both between group patterns and within group patterns. A vector of periodicity plot is useful for determining whether the seasonal pattern is constant over the time period.

The Cycle Plot and Box-Plot

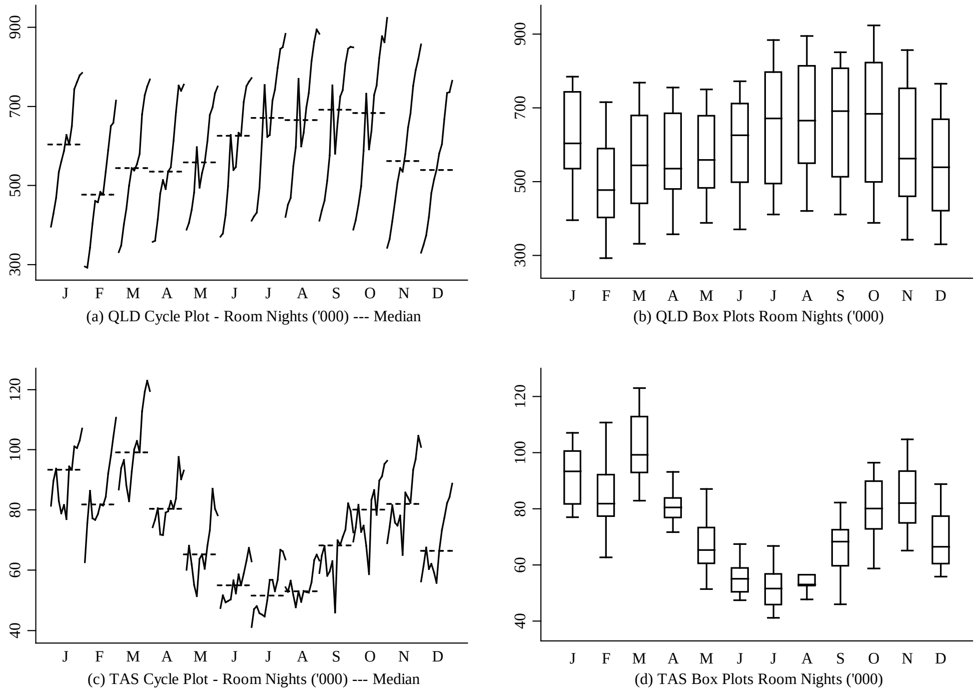
A cycle plot divides the graph into separate plots where each plot contains a time series corresponding to the periodicity of the data. For monthly data for example, there would be 12 plots, the first corresponding to a time series for January, the second for February etc. Figures 23a and 23c are the cycle plots corresponding to monthly data from July 1983 to March 1997 on room nights occupied in hotels and Queensland and Tasmania[19]. The cycle plots highlight the differences in seasonality patterns between the two series. Queensland temperatures throughout the calendar year are generally moderate and the median room nights tend to be higher during the winter months of Australia. Tasmania, on the other hand, can be cold particularly during the winter months and consequently the median room nights drop over this period.

---

[19] Data source: ABS 8635.0.

A series of box plots corresponding to the periodicity of the data can also be used to indicate seasonality as the cycle plot. The box plot provides a summary display of the distribution of the data by graphically showing quartiles of the data (see Hirschberg et al 2005). It shows the median, the spread of the bulk of the data (the length of the box which is the distance from the 25th percentile to the 75th percentile), and how stretched the tails of the distribution are (the length of the lines relative to the box). Additional values that are greater than the limits are sometimes plotted as well. Figures 23b and 23d plot the room nights data using a series of month specific box plots. The seasonal pattern described above is also clearly seen from these sequence of box plots.

Figure 23: Cycle and Box Plots



Vector of Periodicity Plots

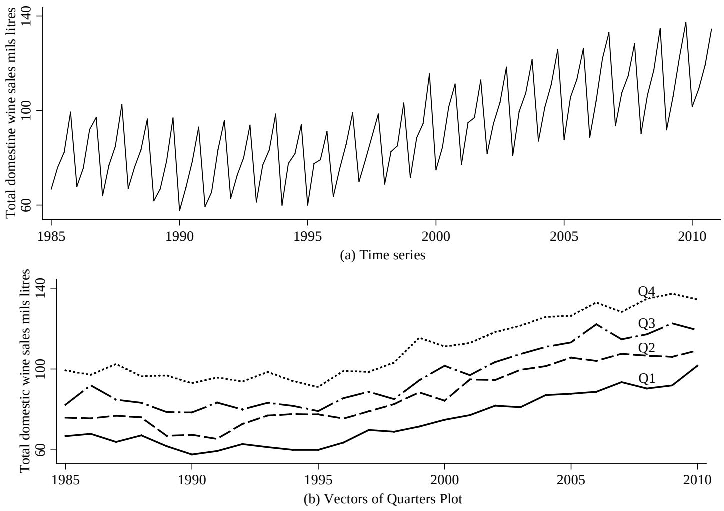
A vector of periodicity plot overlays the time series for each periodicity in a single graph. If the seasonal pattern is constant across the time period, the individual lines should be roughly parallel. This plot is useful in identifying variations in the seasonality over a number of years. In the case of no seasonal patterns the lines will intersect. A varying seasonal pattern is indicated by only some lines intersecting.

Figure 24a is a vector of periodicity plots for quarterly domestic wine sales in Australia from 1985 until 2010[20]. The absence of intersecting lines indicates the presence of a strong seasonal pattern. Figure 24b plots each quarter separately over

[20] Data source: ABS 8504.0.

time. The lines corresponding to each quarter do not intersect which implies all quarters have distinct seasonal patterns that have remained constant over the entire time period.

Figure 24: Seasonality and Vectors of Quarters



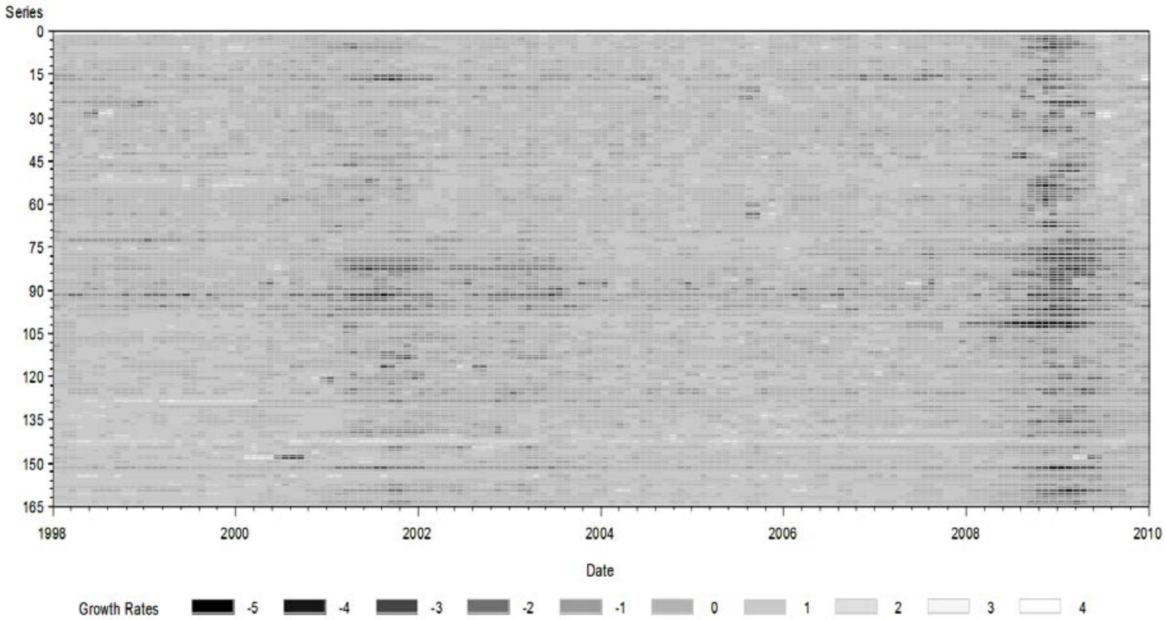
9. Heat Map

A choropleth map with shading is often used to show a theme in cross-section data which is related to a specific geographical location. The first choropleth maps were constructed by Baron Charles Dupin in 1826 who used continuous shading from light to dark to show the geographic distribution of regional values in literacy in France (see Friendly 2008). A Similar map is a heat map which is a graphical representation of data where the values taken by a variable in a two-dimensional table are represented by colours. They have been proposed as a method for examining patterns in large scale data sets.

Time Series Heat Maps

In Figure 25 we demonstrate the construction of a heat map where one dimension is time measured as the month when the observation was made and the other is the series number. The data shown are monthly series for the United States for the period from January 1959 to September 2010 (see Stock and Watson 2014). We plot the monthly growth rates for industrial production classified by industry and market, divided by their standard deviation for the first 164 series from 1998 to 2010. Note that the economic downturns in 2001 and 2008/9 are apparent as darker periods across all sectors. Also, certain sectors such as the Apparel Industry (series 91) show prolonged periods of negative growth.

Figure 25: Heat map of monthly growth rates for industrial production classified by industry and market divided by the series standard deviation



Calendar Plots

A calendar plot is another example of a heat map of daily data. For a single year there are 7 rows that represent days of the week and 53 columns that represent each week in the year. Each day is shaded to reflect the value of the variable for that day which allows daily, weekly, monthly and holiday patterns to be highlighted (see Wicklin 2009). Multiple plots corresponding to a number of years can be used to show how these patterns change over time.

Regular weekly price cycles in retail petrol prices have been observed in major cities in Australia. Vector of periodicity plots of the days of the week prices for 2009 are shown in Figure 26a and for 2010 in Figure 26b[21]. In 2009, prices were lowest on Wednesday and consistently higher on Fridays as shown by the lack of line intersections. However, there was no such consistent pattern in 2010 with both the low and high prices changing across days throughout the year as shown by multiple intersections.

Figure 27 is the equivalent calendar plot of the same data in which prices are shaded different colours from lowest light shades to more expensive dark shades. Over this period, we see a general shift to darker shades which highlights a general increase in prices. For 2009, Wednesdays are associated with lighter shades indicating that it was the least expensive day to buy. We can see towards the end of the week prices increasing with Friday being consistently the most expensive. For 2010 no such consistent pattern appears. In January, Wednesday begins as the least expensive day but by April, May and June we see a move to Friday while during this period Sunday

---

[21] See Footnote 5.

becomes the more expensive day. One advantage of the calendar plot is the ability to locate particular days such as holidays in the data.

Figure 26: Daily Retail Petrol Prices

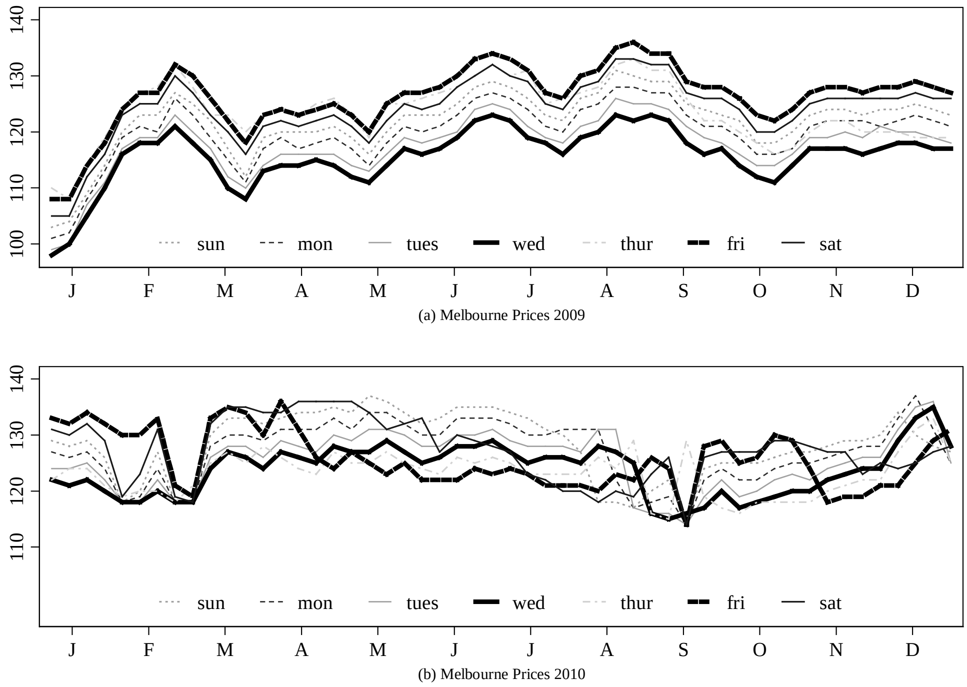
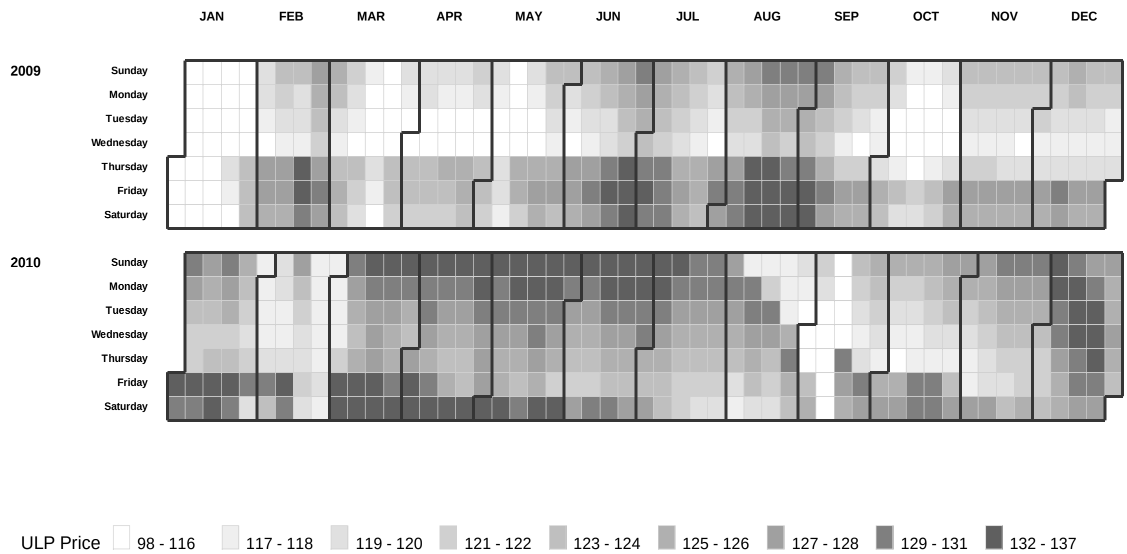


Figure 27: Calendar Plot of Daily Retail Petrol Prices



10. Conclusions

In this paper we have illustrated numerous types of plots for the presentation of time series data. In the construction of any plot it is important to remember that the type of plot should be chosen to emphasize important patterns in the data.

Furthermore, all plots should be legible and well designed. The visual impact of the plot should be maximised by changing the default options in the computer software used. Do not clutter the plot by using unnecessary visual elements.

Adding colour to a plot can be a useful way of drawing attention to particular features. Examples include the use of bright or dark colours such as red and black to emphasize the important line or for example encoding different categories in a scatter plot. However, many journals print graphs in black and white with the colour version appearing online. Remember to print the graph in black and white or make a copy in gray scale to test whether the contrast in values and line types is sufficient to distinguish differences.

There are numerous resources for producing plots in computer programs. For Stata, a useful resource is Mitchell (2012), for R is Wickham (2016), for SAS is Liu (2015), for Matlab is Holland (2017) and Eviews has a basic graphic tutorial available at https://www.eviews.com/Learning/graphs.html.

References

AGS (2019), Australian Gambling Statistics 1992–93 to 2017–18 35th edition, Queensland Government Statistician's Office, Queensland Treasury.

Barnes, T., M. Stevens, M., Thoss and A. Taylor (2017), *The size and growth of the Northern Territory's Gambling Industry*, Northern Institute Working Paper 08/2017, Northern Institute and Menzies School of Health Research.

Carson, C. (2009), "Numbers as pictures: Examples of data visualization from the Business Employment Dynamics program", *Section on Government Statistics, Joint Statistical Meetings* downloaded w*ww.amstat.org/sections/srms/proceedings/y2009/Files/304554.pdf*.

Cleveland, W. (1985), *The elements of graphing data*, AT&T Bell Laboratories, N.J.

Costigan-Eaves P. and M. Macdonald-Ross (1990), "William Playfair (1759-1823)", *Statistical Science*, 5, 318-326.

Demir I. and R. D. Tollison, (2015) "Graphs in Economics", *Economics Bulletin*, 35, 1834-1847.

Finkelstein, M. and B. Levin (2001), *Statistics for Lawyers*, 2nd ed. Springer-Verlag: New York.

Friendly, M. (2008), "The Golden Age of Statistical Graphics", *Statistical Science*, 23, 502–535.

Holland O. (2017), *Graphics and GUIs with MATLAB*, 3rd edition, Chapman and Hall/CRC.

Hirschberg, J., L. Lu and J. Lye (2005), 'Descriptive methods for cross-section data', *Australian Economic Review*, 38, 333-50.

Liu C. (2015), Producing High-Quality Figures Using SAS/GRAPH® and ODS Graphics Procedures, Taylor & Francis Group.

Martino, J. (1971), "Examples of Technological Trend Forecasting for Research and Development Planning", *Technological Forecasting and Social Change*, 2, 247-260.

Milroy, R. and E. Poulten (1978), "Labelling Graphs for Improved Reading Speed", *Ergonomics*, 21, 55-61.

Mitchell, M. (2012), *A Visual Guide to Stata Graphics*, 3rd edition, Stata Press, Texas US.

Pole, A., M. West and J. Harrison (1994) *Applied Bayesian Forecasting and Time Series Analysis*, Chapter 6, Chapman and Hall New York.
Stock, J. and M. Watson (2014), "Estimating Turning Points using large data sets", *Journal of Econometrics*, 178, pp. 368-381.
Tufte, E (2006), *Beautiful evidence*, Graphics Press.
Wei, W. (2006), *Time Series Analysis Univariate and Multivariate Methods*, 2nd edition, Pearson Education Inc.
Wickham, H. (2016), *ggplot2: Elegant Graphics for Data Analysis* 2nd edition, Springer
Wicklin, R. (2009), "High-Flying Graphics at the 2009 Data Expo", *Statistical Computing and Graphics*, 20, 5-11.