



ECOM5005 Tutorial 2 - Detailed Solutions

Topic 2: Organising and Presenting Data

Curtin University

School of Accounting, Economics and Finance

Aligned with Topic 2: Organising and Presenting Data



How to use this solution

Follow the tasks in sequence. The Part 2 instructions assume very limited Excel experience and explain the required clicks, formulas, error checks and expected results. Keep the supplied source worksheets unchanged. The charts shown are the final outputs students should reproduce in Excel; minor differences in font or colour are acceptable, but titles, variables, dates, scales and numerical results must be correct.

Part 1 - Detailed solutions

Question 1. Stem-and-leaf display

The stem gives the tens digit and each leaf gives the units digit. The complete ordered data are:

9, 13, 16, 16, 18, 21, 22, 23, 25, 26, 29, 30, 41, 42, 43, 43, 45, 47.

“No more than 30” means less than or equal to 30, so the values included are:

9, 13, 16, 16, 18, 21, 22, 23, 25, 26, 29, 30.

There are **12 stores** with no more than 30 pairs of glasses.

A stem-and-leaf display preserves every original value. Always use the key to identify place value, include repeated leaves separately, and translate verbal boundaries such as “no more than” into mathematical inequalities.

Question 2. Contingency-table interpretation

The correct answer is D. Each statement is checked using the relevant conditional group.

Option	Evaluation	Conclusion
A	Among the 38 students with allergies, Winter has 5, Spring 6, Summer 13 and Fall 14. Winter, not Spring, is least frequent.	False
B	Among students without allergies, Summer is most frequent ($23 > 16 > 9 > 7$). Among students with allergies, Fall is most frequent ($14 > 13 > 6 > 5$).	False
C	Among the 55 students without allergies, Spring has 7 students and Fall has 16 students.	False
D	In every season, the no-allergy count exceeds the allergy count: $5 < 9$, $6 < 7$, $13 < 23$ and $14 < 16$.	True

The row percentages also confirm the final statement:

Season	Has allergies	No allergies
Winter	35.71%	64.29%
Spring	46.15%	53.85%
Summer	36.11%	63.89%
Fall	46.67%	53.33%

D is true: for every favourite-season category, a surveyed student was more likely to have no allergies than to have allergies.

Row percentages answer questions conditional on the row category; column percentages answer questions conditional on the column category. A correct numerator divided by the wrong denominator produces an incorrect interpretation.

Question 3. Comparing speed distributions

Using Excel-style inclusive quartiles, the numerical summaries are:

Measure	Northbound	Southbound
Count	19	20
Mean	65.526	63.050
Minimum	60.0	55.0
First quartile Q1	63.0	59.5
Median	65.0	63.5
Third quartile Q3	66.0	67.25
Maximum	83.0	71.0
Interquartile range	3.0	7.75
Sample standard deviation	4.812	4.774
Upper outlier fence	70.5	78.875

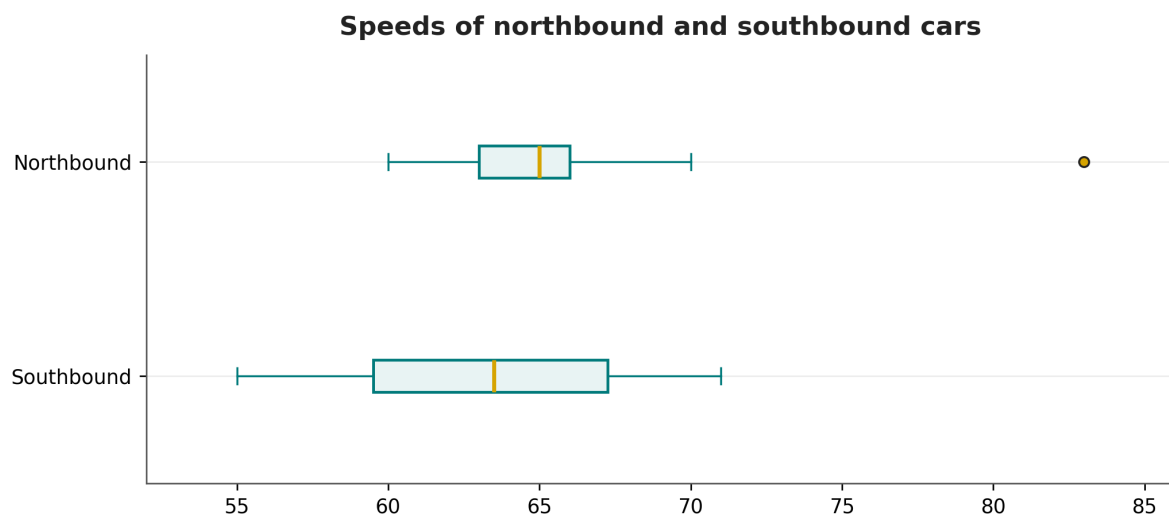


Figure 1. Final comparison box plots. The northbound value of 83 mph is plotted as an outlier

Interpretation

- Centre: the medians are similar, although the northbound median is 1.5 mph higher (65.0 versus 63.5).
- Middle spread: southbound speeds are much more dispersed in the middle 50% because the southbound IQR is 7.75 mph, compared with 3.0 mph northbound.
- Unusual observation: 83 mph exceeds the northbound upper fence of 70.5 mph and is therefore an outlier under the $1.5 \times \text{IQR}$ rule.
- Overall standard deviation: the two sample standard deviations are similar because the northbound outlier offsets the otherwise narrow northbound central box.
- Scope: the data describe one location and one 15-minute period. They do not establish a stable directional difference for the highway generally.

Excel steps: Creating the box plots

Enter the two series in separate Excel columns with headers. Select both columns, then choose Insert > Insert Statistic Chart > Box and Whisker. Use Chart Design > Switch Row/Column if the groups appear incorrectly. Add the title and horizontal-axis label, and ensure the outlier marker is visible.



Question 4. Airport baggage delays

The observed counts are:

Itinerary	Not delayed	Delayed	Total
Direct	550	50	600
One connection	200	40	240
Two or more	60	30	90
Total	810	120	930

(a) Percentage of the overall total

Every cell is divided by 930.

Itinerary	Not delayed	Delayed	Total
Direct	59.14%	5.38%	64.52%
One connection	21.51%	4.30%	25.81%
Two or more	6.45%	3.23%	9.68%
Total	87.10%	12.90%	100.00%

(b) Delayed-bag rate within each itinerary type

Each delayed count is divided by its row total:

- Direct: $50 / 600 = 8.33\%$.
- One connection: $40 / 240 = 16.67\%$.
- Two or more connections: $30 / 90 = 33.33\%$.

The observed delay rate rises sharply with the number of connections.

(c) Itinerary composition among delayed bags

The denominator is now the 120 delayed bags:

- Direct: $50 / 120 = 41.67\%$.
- One connection: $40 / 120 = 33.33\%$.
- Two or more connections: $30 / 120 = 25.00\%$.

(d) Evaluating the manager's claim

The claim is misleading. Direct itineraries generate the largest count of delayed bags because they account for 600 of 930 audited bags. Risk is better represented by the within-itinerary delay rate. On that measure, direct flights have the lowest observed risk (8.33%) and itineraries with two or more connections have the highest (33.33%).

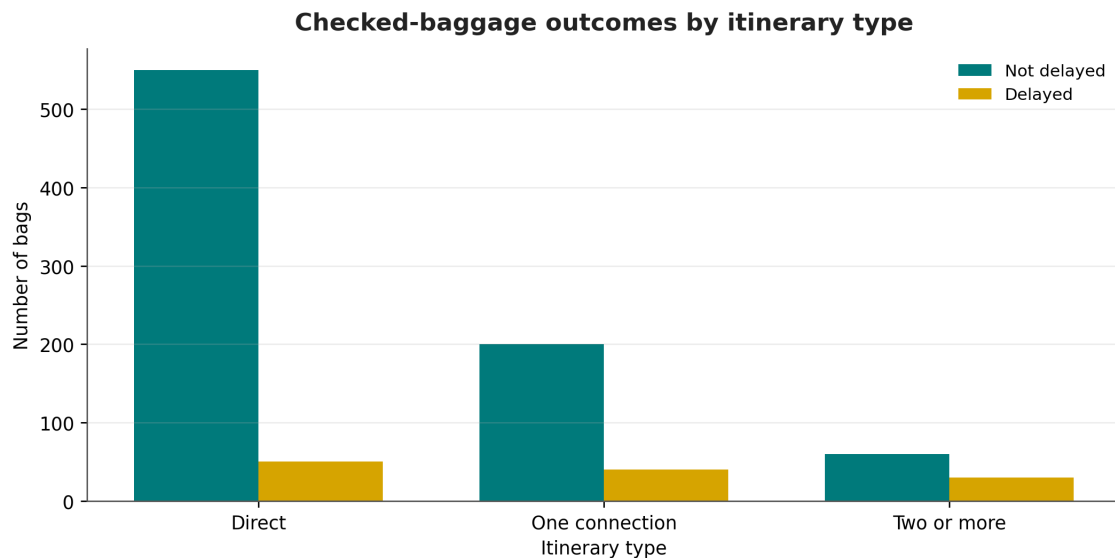


Figure 2. A count chart is useful for workload, but row percentages are required for comparing risk

(e) Recommended communication

Use a 100% stacked bar chart or a clustered column chart of delay percentages within itinerary type. This puts all three itineraries on the same 0–100% basis and directly answers the risk question. A count chart can be shown separately for operational workload.

(f) Limitation

This is an observational table. It does not control for airline, airport, weather, baggage-handling contractor, connection time or passenger volume. Connections may be associated with delays without being the sole cause.

Counts describe workload and total impact. Conditional percentages describe rate or risk. Management often needs both, but they must not be substituted for one another.

Question 5. Electricity demand and temperature

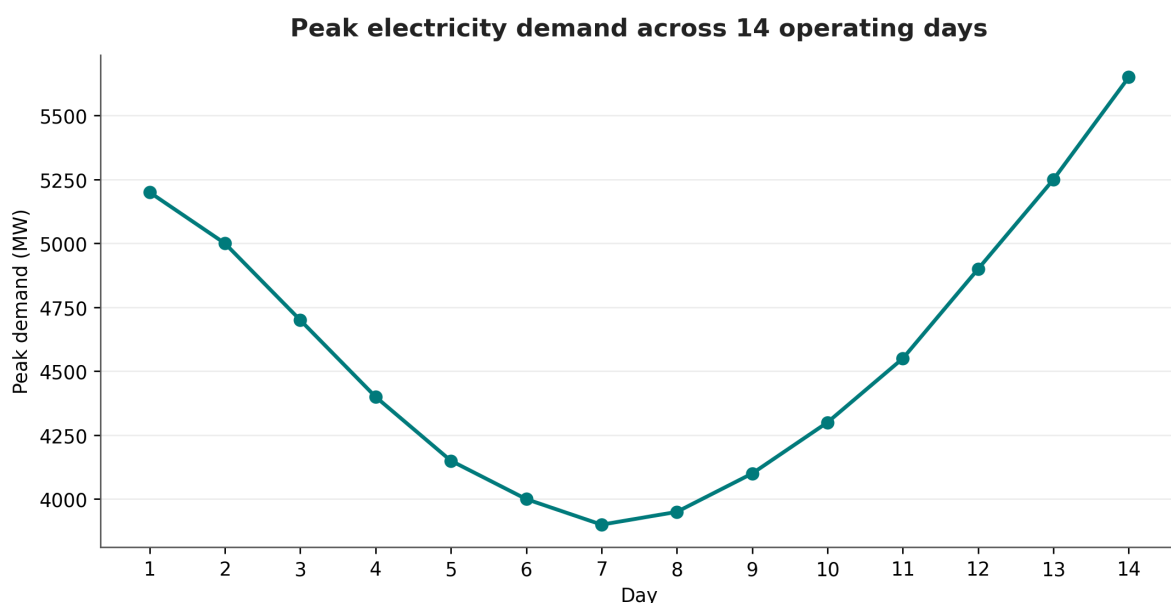


Figure 3. Time-series plot of peak demand by operating day

The time-series plot shows demand falling from 5,200 MW on Day 1 to 3,900 MW on Day 7, then increasing to 5,650 MW on Day 14. It identifies the timing and sequence of changes.

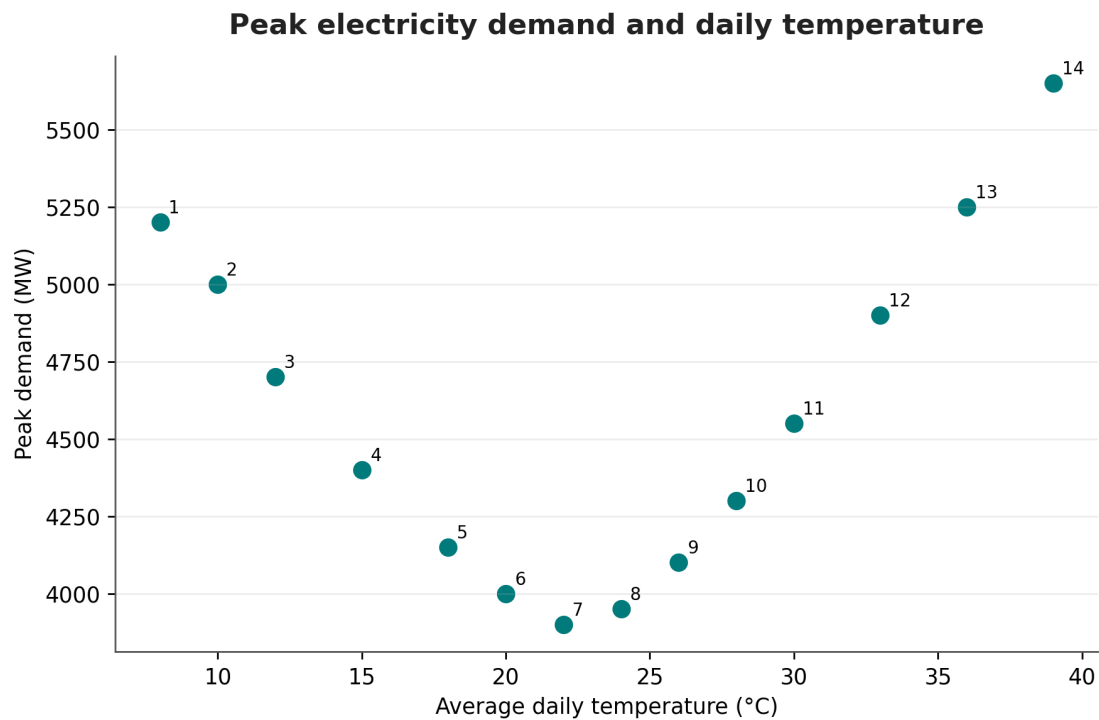


Figure 4. Scatter diagram of temperature and peak demand. Day labels preserve operational context

The scatter diagram shows a pronounced curved, approximately U-shaped relationship: demand is high on very cold and very hot days and lowest around mild temperatures. A single straight line would average across the two arms of the curve and could obscure the operationally important pattern.

Summary	Result
Average peak demand	4,575 MW
Minimum	3,900 MW on Day 7
Maximum	5,650 MW on Day 14
Range	1,750 MW

A vertical axis beginning at 3,850 MW would exaggerate changes in bar height or line movement. Three-dimensional effects add no information and can distort perceived distance. A defensible redesign uses a flat two-dimensional chart, a clearly labelled axis and an honest scale. For a line chart, starting at zero is preferred when the visual comparison depends on magnitude; if a non-zero scale is justified to show small temporal movements, the truncation must be obvious and explained.

Potential omitted variables include humidity, day of week, industrial load, public holidays, solar generation, electricity price and conservation alerts. The graphs show association and sequence, not a controlled causal effect.



Part 2 - Excel solution

Dataset at a glance

The supplied UE_Data worksheet occupies A1:D510. It contains 509 visible monthly observations from February 1978 to June 2020. The columns are Time, Persons, Males and Females. The three unemployment-rate series contain no blank cells. The rates are percentages and the Metadata worksheet describes the series as monthly, original percentage data.

Before beginning

- Use desktop Microsoft Excel. The web version may display chart and formatting options differently.
- Preserve UE_Data and Metadata. Work on copies or new worksheets.
- Do not delete unusual values simply because a peak looks large. Check the source and metadata first.
- Save frequently as an Excel workbook, not CSV, because CSV cannot preserve worksheets, formulas or charts.

Task A - Save and audit the workbook

Excel steps: Step 1: save a working copy

1. Open unemp_AUS.xlsx in desktop Excel.
2. Choose File > Save As.
3. Name the file Tutorial 2 - Your Name.xlsx.
4. Confirm the format is Excel Workbook (*.xlsx) and click Save.
5. Click the plus sign twice and rename the new sheets Analysis and Dashboard.

Excel steps: Step 2: build a dataset-audit table

On Analysis, enter headings in A1:C1: Check, Formula, Expected result. Enter the formulas shown in the following table.

Check	Formula	Expected
Visible dates	=COUNTA(UE_Data!A2:A510)	509
Numeric Persons rates	=COUNT(UE_Data!B2:B510)	509
Numeric Male rates	=COUNT(UE_Data!C2:C510)	509
Numeric Female rates	=COUNT(UE_Data!D2:D510)	509
Blank Persons cells	=COUNTBLANK(UE_Data!B2:B510)	0
Blank Male cells	=COUNTBLANK(UE_Data!C2:C510)	0
Blank Female cells	=COUNTBLANK(UE_Data!D2:D510)	0
First visible date	=MIN(UE_Data!A2:A510)	Feb-1978
Last visible date	=MAX(UE_Data!A2:A510)	Jun-2020

Format the date-result cells using Home > Number Format > Custom and type mmm-yyyy.

Checkpoint: audit results

The visible workbook has 509 monthly records, not 516. The Metadata worksheet reports “No. Obs” as 516. This is a genuine reconciliation issue: report it and use the actual visible rows for the analysis. Do not invent or fill seven unobserved months.



Metadata description

Series	Unit/frequency	Type	Series ID
Persons	Percent; Month	Original	A84423092X
Males	Percent; Month	Original	A84423078C
Females	Percent; Month	Original	A84423106A

Task B - Prepare the Analysis worksheet

Excel steps: Copy the data and add helper columns

1. On UE_Data, click cell A1.
2. Press Ctrl+A once, or select A1:D510 using the Name Box.
3. Press Ctrl+C.
4. Go to Analysis, click A12 and press Ctrl+V.
5. The copied data should occupy A12:D521. This keeps the audit table above the data.

Because the copied data start at row 12, the first observation is row 13. Add these helper columns:

Cell	Heading/formula	Purpose
E12	Year	Calendar year
F12	Month	Readable month label
G12	Female - Male gap	Percentage-point difference
E13	=YEAR(A13)	Extract year
F13	=TEXT(A13,"mmm")	Extract abbreviated month
G13	=D13-C13	Female rate minus male rate

Select E13:G13, then double-click the fill handle to copy the formulas to row 521.

Checkpoint

For the final row, June 2020, Year should be 2020, Month should be Jun and the gap should be approximately -0.107 percentage points.

Task C - Create the four required line charts

The original worksheet rows for the required start dates are:

Start date	UE_Data row	Number of months to Jun-2020
Jan-1980	25	486
Jan-2000	265	246
Jan-2010	385	126
Jan-2020	505	6
Jun-2020	510	—

Excel steps: Persons chart for January 2010 to June 2020

1. Go to UE_Data.
2. Select A385:B510. Column A contains dates and column B contains Persons.
3. Choose Insert > Line or Area Chart > 2-D Line.
4. If the legend says "Series 1", right-click the chart, choose Select Data, edit the series name and type Persons.
5. Set the chart title to Australia unemployment rate - Persons, Jan 2010 to Jun 2020.
6. Add the vertical-axis title Unemployment rate (%) and horizontal-axis title Month.
7. Right-click the vertical axis, choose Format Axis, set Minimum to 0 and Maximum to 8.
8. Right-click the horizontal axis and select a readable date unit. Labels every 12 or 24 months are usually sufficient.

Repeat the procedure with A265:B510 for 2000-2020 and A25:B510 for 1980-2020. Use vertical maxima of approximately 9 and 13 respectively.

Excel steps: Male-Female comparison chart

1. Select A25:A510.
2. Hold Ctrl on Windows or Command on Mac and select C25:D510.
3. Insert a 2-D line chart.
4. Confirm that the two series are named Males and Females and the dates appear on the horizontal axis.
5. Use the title Australia unemployment rate by recorded sex, Jan 1980 to Jun 2020.
6. Set the vertical axis from 0 to 14 and retain a clear two-series legend.

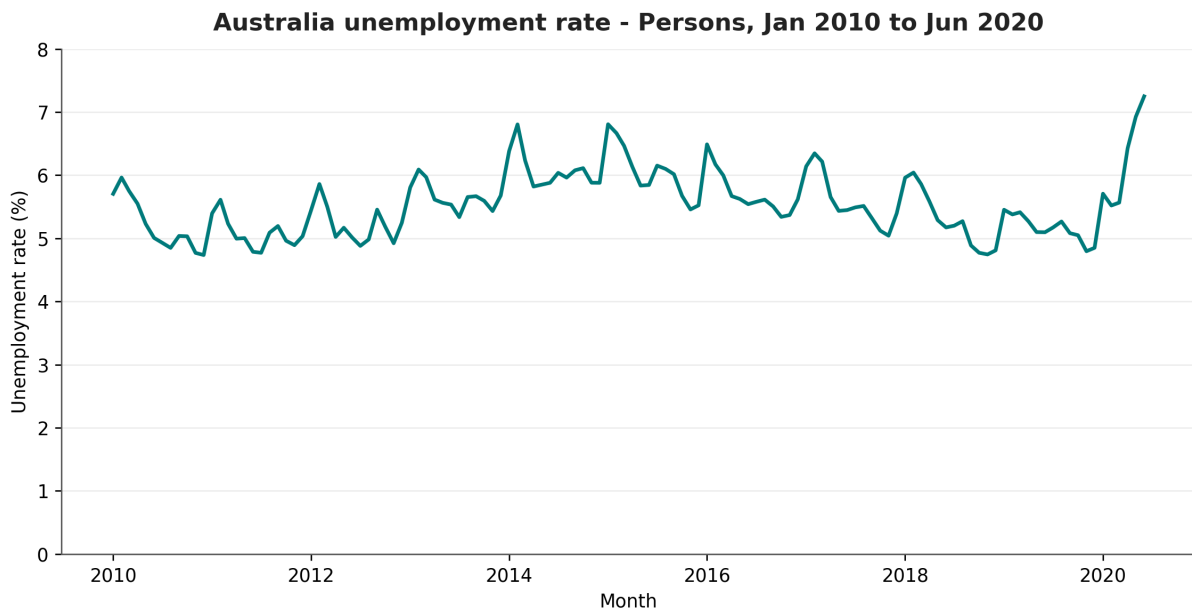


Figure 5. Expected final Persons chart for January 2010 to June 2020

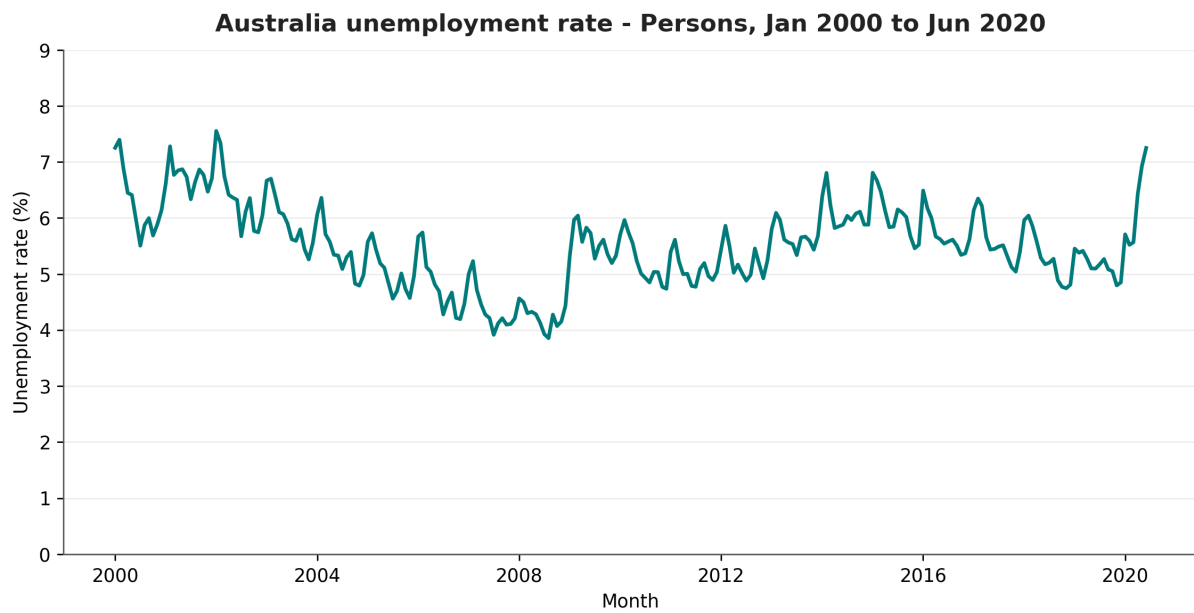


Figure 6. Expected final Persons chart for January 2000 to June 2020

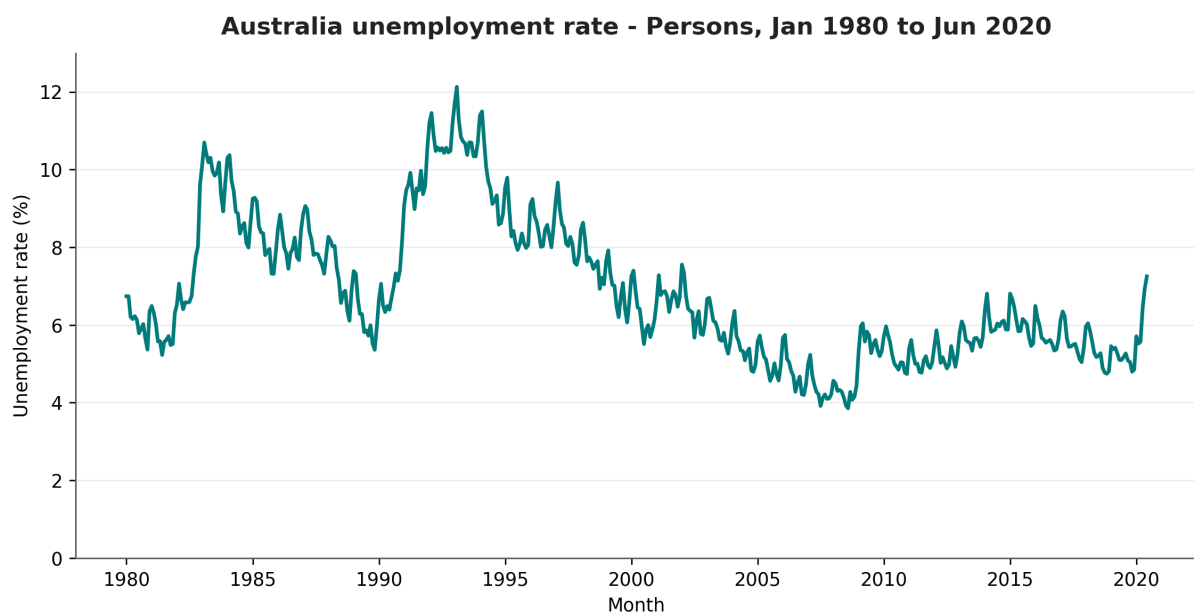


Figure 7. Expected final Persons chart for January 1980 to June 2020

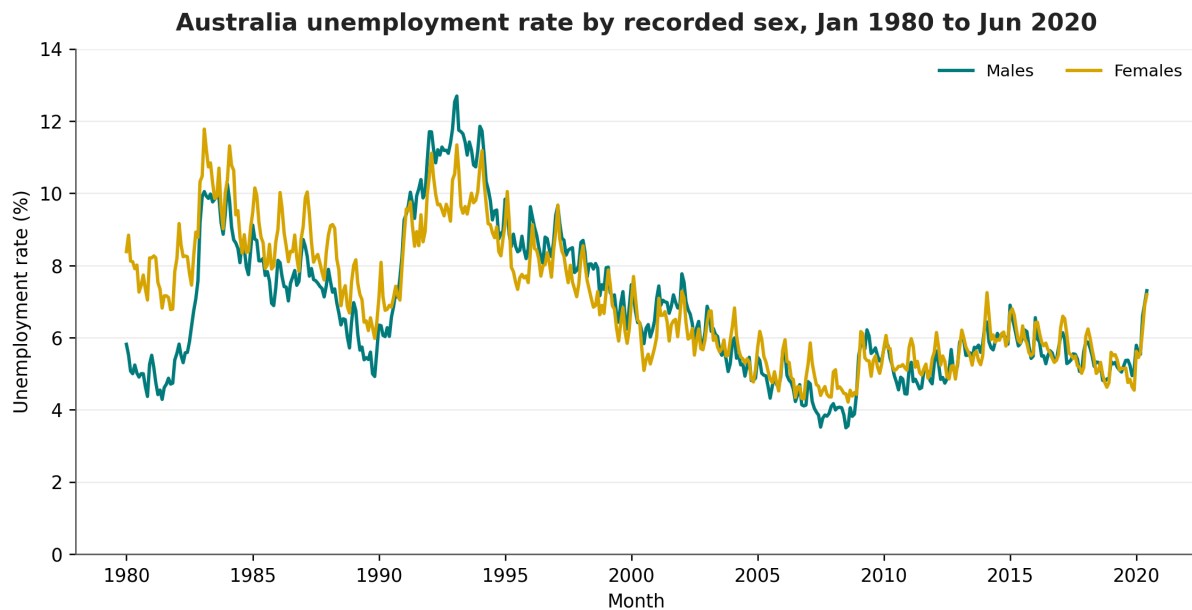


Figure 8. Expected final Male-Female comparison chart

Different vertical-axis maxima change the apparent steepness of movements. Different scales are reasonable for charts covering different horizons, but the scale must be clear. When comparing Male and Female lines on one chart, both series must share the same axis.

Task D - Summary tables

The formulas below use the original UE_Data row ranges. For the 2010 window, enter:

Measure	Formula	Result
Count	=COUNT(UE_Data!B385:B510)	126
Average	=AVERAGE(UE_Data!B385:B510)	5.542
Minimum	=MIN(UE_Data!B385:B510)	4.736
Minimum date	=INDEX(UE_Data!A385:A510,MATCH(MIN(UE_Data!B385:B510),UE_Data!B385:B510,0))	Dec-2010
Maximum	=MAX(UE_Data!B385:B510)	7.250
Maximum date	=INDEX(UE_Data!A385:A510,MATCH(MAX(UE_Data!B385:B510),UE_Data!B385:B510,0))	Jun-2020
Starting rate	=UE_Data!B385	5.706
Ending rate	=UE_Data!B510	7.250
Change	=UE_Data!B510-UE_Data!B385	+1.544

Repeat the formulas using rows 265:510 and 25:510. The final window comparison is:

Window	n	Avg.	Min.	Min date	Max.	Max date	Start	End/change
2010-Jun 2020	126	5.542	4.736	Dec-2010	7.250	Jun-2020	5.706	7.250 / +1.544
2000-Jun 2020	246	5.508	3.854	Aug-2008	7.553	Jan-2002	7.253	7.250 / -0.003
1980-Jun 2020	486	6.827	3.854	Aug-2008	12.127	Feb-1993	6.738	7.250 / +0.512

Decade-style summary

Period	n	Average	Minimum	Maximum	Peak date
1980s	120	7.586	5.226	10.694	Feb-1983
1990s	120	8.772	6.064	12.127	Feb-1993
2000s	120	5.471	3.854	7.553	Jan-2002
2010s	120	5.508	4.736	6.807	Jan-2015
Jan-Jun 2020	6	6.233	5.521	7.250	Jun-2020

Task E - Gender differences and 2020 movement

On the copied Analysis data, the January 1980 observation is row 36 and June 2020 is row 521 because the data begin at row 12. If using column G for Female minus Male:

Measure	Formula	Result
Average gap	=AVERAGE(G36:G521)	+0.236 pp
Largest positive gap	=MAX(G36:G521)	+3.456 pp
Date of largest positive gap	=INDEX(A36:A521,MATCH(MAX(G36:G521),G36:G521,0))	Mar-1981
Largest negative gap	=MIN(G36:G521)	-2.199 pp
Date of largest negative gap	=INDEX(A36:A521,MATCH(MIN(G36:G521),G36:G521,0))	May-1993

Female minus male unemployment-rate gap, Jan 1980 to Jun 2020

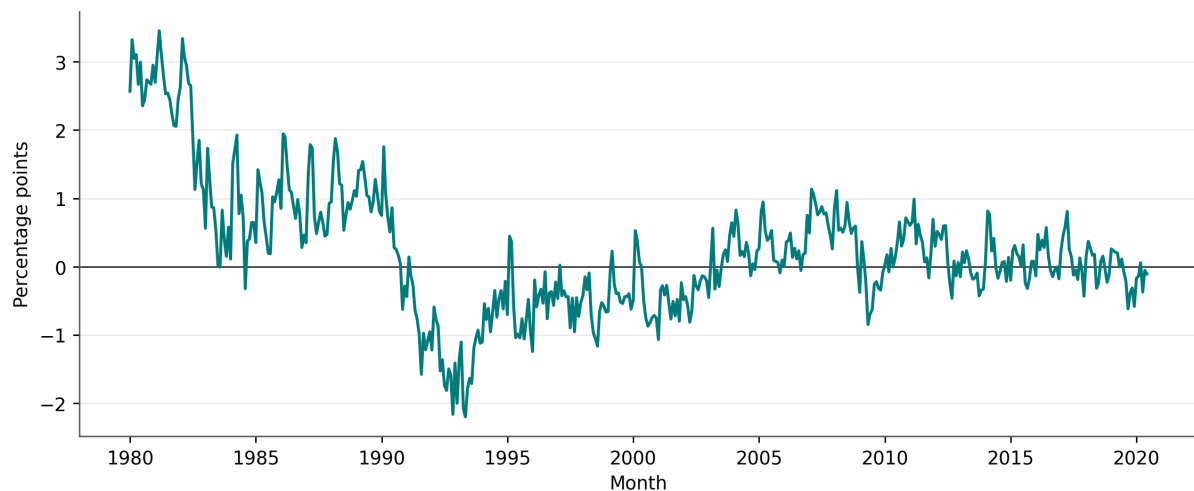


Figure 9. Expected gender-gap chart. Values above zero mean the Female rate exceeds the Male rate

January to June 2020

Series	Jan-2020	Jun-2020	Change (pp)
Persons	5.707	7.250	+1.543
Males	5.788	7.300	+1.513
Females	5.618	7.193	+1.575

The rate increased markedly for all three series. In this six-month interval the Female increase was slightly larger than the Male increase, although the June Male rate remained slightly higher.

Task F - Build the Dashboard

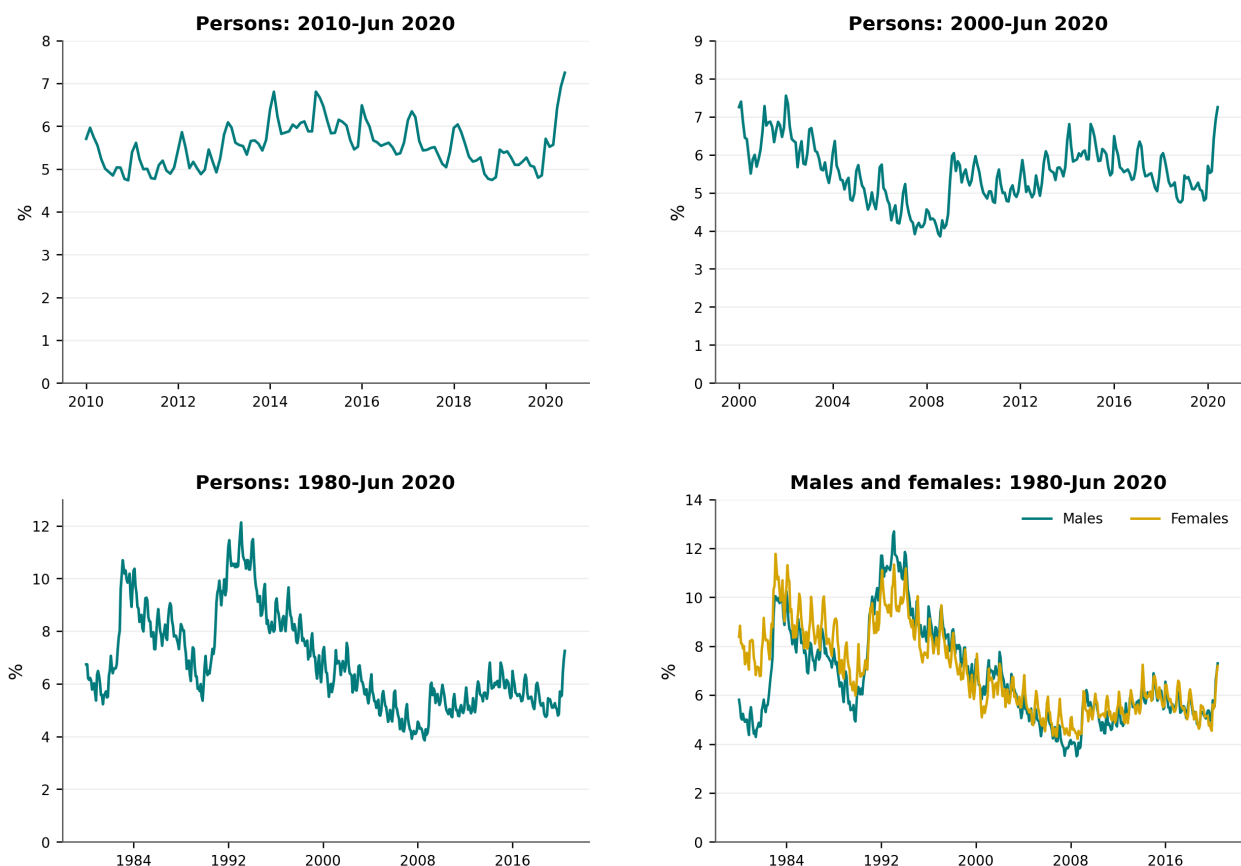
Excel steps: Assemble the dashboard

1. Go to Dashboard and type Australia unemployment-rate dashboard in A1.
2. Move or copy the four charts to the Dashboard. Arrange them in a two-by-two grid.
3. Use Chart Format > Align to align chart edges and Distribute Horizontally/Vertically to create even spacing.
4. Create a KPI block in cells A30:H34. Use formulas linked to the source data rather than typing values manually.
5. Add the source note: Source: supplied ABS-derived workbook unemp_AUS.xlsx; monthly original series.
6. Remove unnecessary chart borders, backgrounds and repeated legends. Keep axes and units clear.

Recommended KPI values are:

KPI	Value
Persons rate, Jun-2020	7.250%
Persons change, Jan-Jun 2020	+1.543 percentage points
Maximum Persons rate, Jan-1980-Jun-2020	12.127% in Feb-1993
Minimum Persons rate, Jan-1980-Jun-2020	3.854% in Aug-2008

Australia unemployment-rate dashboard



Source: supplied ABS-derived workbook unemp_AUS.xlsx; monthly original series.

Figure 10. Example of the completed dashboard layout



Task G - Model written interpretation

Model interpretation

The 1980-2020 chart reveals substantial long-run variation that is hidden in the shorter charts. The Persons unemployment rate rose to approximately 10.69% in February 1983 and reached the sample-window maximum of 12.13% in February 1993. It then generally declined, reaching a minimum of 3.85% in August 2008. The 2000-2020 view highlights a relatively high rate in January 2002, the low point in 2008 and a gradual increase into the middle of the 2010s. By contrast, the 2010-2020 chart places much greater visual emphasis on the sharp movement during the first half of 2020. The Persons rate increased from 5.71% in January to 7.25% in June, a rise of 1.54 percentage points.

The gender comparison shows that the Female rate was often higher in the early part of the sample, but the gap changed sign. The average Female-minus-Male gap from January 1980 to June 2020 is about +0.24 percentage points. The largest positive gap is +3.46 points in March 1981, while the largest negative gap is -2.20 points in May 1993. Therefore, the relative position of the two series is not constant across time.

These figures describe historical association and timing; they do not identify causes. The workbook contains original monthly series and does not control for population composition, participation, industry structure, policy changes or economic shocks. The visible worksheet also contains 509 observations while the metadata reports 516, so the discrepancy should be documented. Conclusions should be confined to the supplied data and supported with source documentation before external use.

Common Excel problems and fixes

Problem	Likely cause	Fix
Dates appear as numbers such as 43831	Date cells use General format	Format as mmm-yyyy or yyyy-mm.
Chart shows row numbers instead of dates	Date range was not selected as categories	Use Select Data and assign column A as horizontal labels.
Chart contains one line per month	Data orientation is wrong	Use Chart Design > Switch Row/Column.
Male and Female lines have separate axes	Secondary axis was added	Plot both on the primary axis.
The source data changed after sorting	Only one column was sorted	Undo and always expand the entire selection.
Metadata count and visible count differ	Workbook is a trimmed extract	Report the discrepancy; analyse visible rows only.
Dashboard is unreadable	Too many labels or chartjunk	Reduce date labels, remove 3-D effects and retain only essential legends.
The line appears exaggerated	Vertical axis is severely truncated	Start at zero or clearly justify and disclose truncation.
Formula shows #N/A	MATCH range does not align with value range	Ensure date and rate ranges start and end on the same rows.

Final submission checklist

1. The workbook is saved as .xlsx.
2. UE_Data and Metadata remain unchanged.
3. Analysis contains the audit, copied data, helper columns and summary tables.
4. The visible count of 509 and metadata count of 516 are both reported.
5. All four required time-series charts have correct date windows, titles and units.
6. The gender-gap chart and January-June 2020 change table are complete.
7. The window and decade summary tables reconcile with the charts.
8. The Dashboard contains four charts, KPI values and a source note.
9. The written interpretation distinguishes description from causation.
10. The workbook contains formulas and editable charts, not screenshots only.



Part 3 - Pearson MyLab

Complete Quiz 1. When seeking technical help, report the platform, browser, course identifier and exact error message so that the issue can be diagnosed efficiently.