Time series.

Using the statistical enquiry cycle to investigate time series data involves:
- using existing data sets
- selecting a variable to investigate
- selecting and using appropriate display(s)
- identifying features in the data and relating this to the context
- finding an appropriate model
- using the model to make a forecast
- communicating findings in a conclusion.

Example:
Alcohol consumption in New Zealand by type of alcohol.
Reference and data description.
SOURCE: http://www.stats.govt.nz/infoshare/#Under Industry Series
TIME FRAME: 1995 to 2012 inclusive

DESCRIPTION: Total quantities (millions of litres) of types of alcohol consumed by quarter
VARIABLES: Litres of Alcohol Consumed (Qrtly-Mar/Jun/Sep/Dec) AC_beer AC_spirits AC_wine

DESCRIPTION: Litres of Alcohol Per Head of Population by quarter
VARIABLES: Litres of Alcohol Per Head of Population (Qrtly-Mar/Jun/Sep/Dec) LPH_15yr_over LPH_18yr_over LPH_20yr_over LPH_Total_Popn

Quarter	AC_beer	AC_spirits	AC_wine	LPH_15yr_over	LPH_18yr_over	LPH_20yr_over	LPH_Total_Popn
1995Q1	3.368	0.939	1.445	2.088	2.213	2.308	1.605
1995Q2	3.235	0.951	1.734	2.162	2.295	2.397	1.659
1995Q3	3.157	1.189	1.67	2.203	2.339	2.441	1.691
1995Q4	4.335	1.15	2.385	2.834	3.004	3.133	2.18
1996Q1	3.373	0.882	1.511	2.036	2.157	2.246	1.57
1996Q2	3.259	1.155	1.74	2.157	2.287	2.382	1.659
1996Q3	2.89	0.871	1.73	1.919	2.034	2.118	1.476

INTRODUCTION .
1) Description & Investigative Question

2) Source of Data
3) Aim / Interest

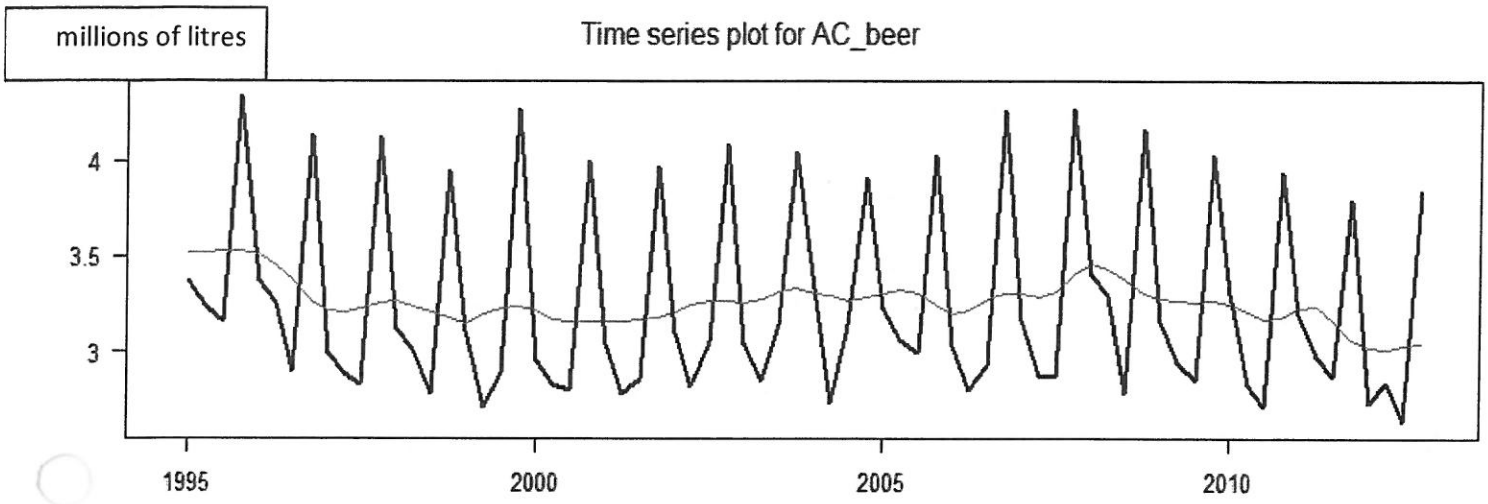
4) Variables Defined & Described

5) Hypothesis 'I think that... because...

6) Research findings summarized

Alcohol consumption in New Zealand by type of alcohol.
THE TREND

1) Graph from iNZight
2) List the 'obvious' points to discuss



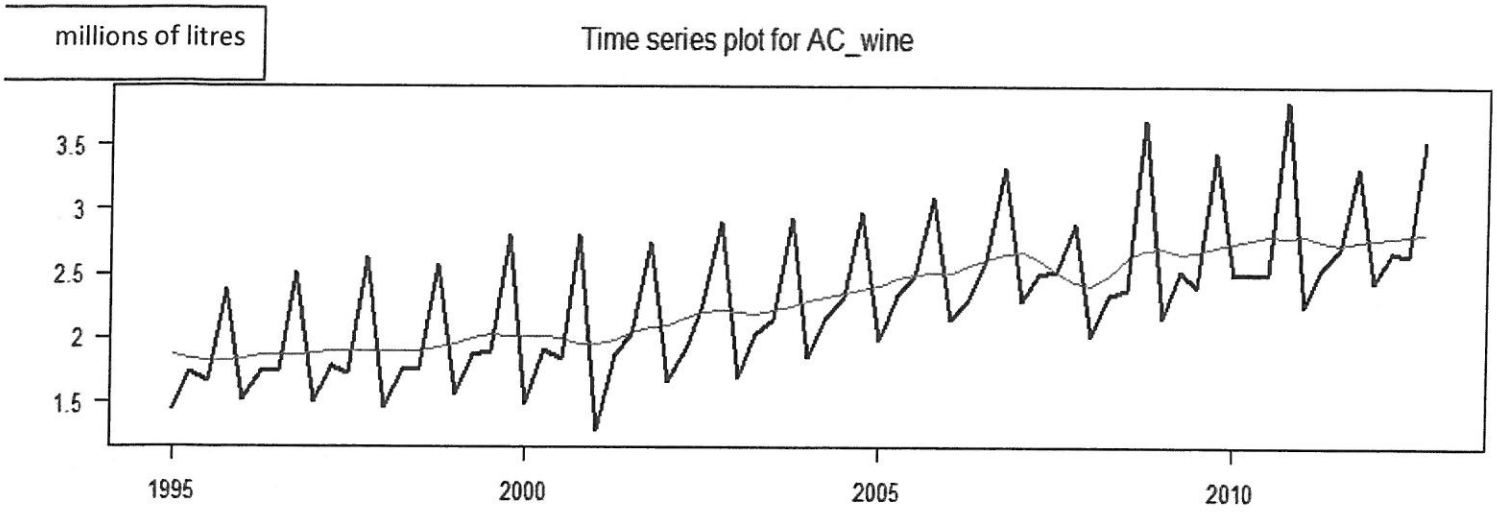
- Obvious (State the obvious)

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)



- Obvious (State the obvious)

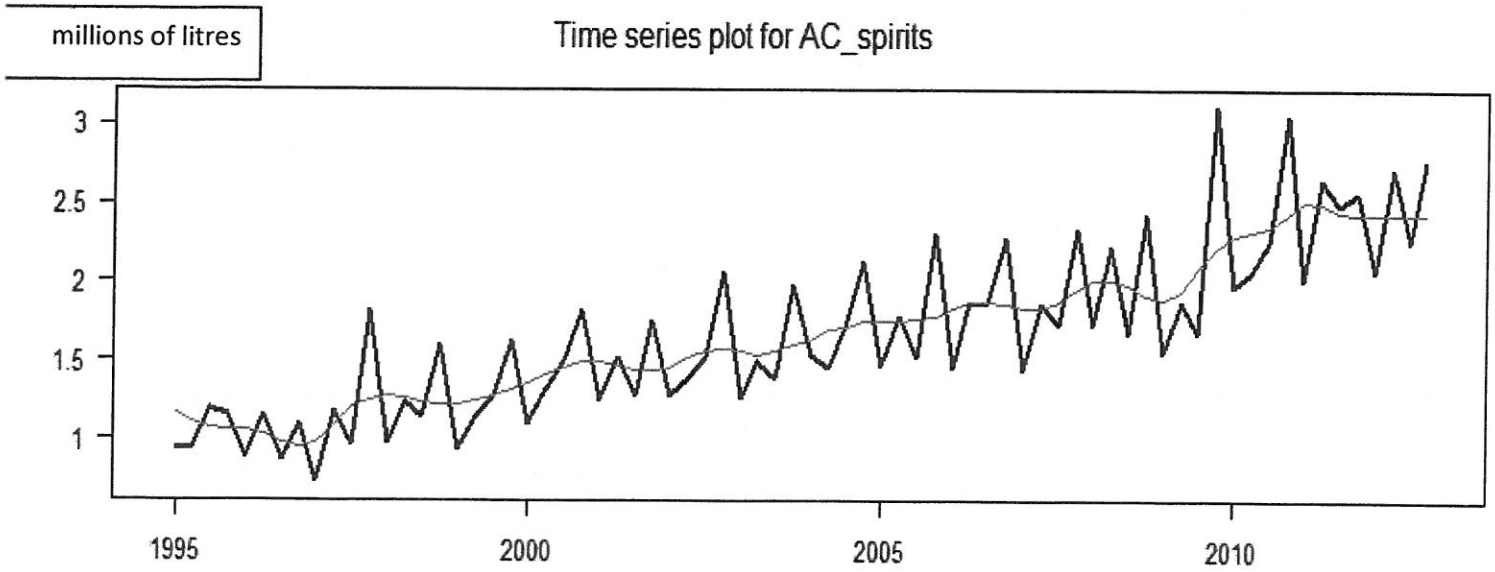
Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

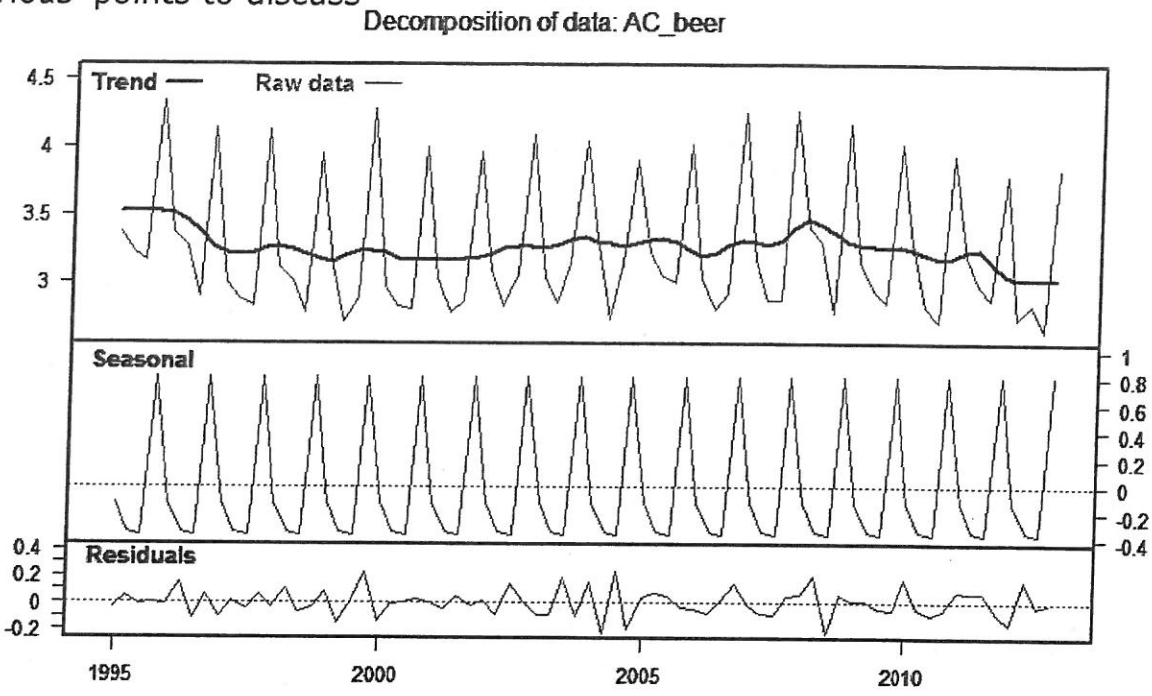
- Relate (To references / research and hypothesis made)

(Do this one on your own paper)



Alcohol consumption in New Zealand by type of alcohol.
THE DECOMPOSED DATA

1) Graph from iNZight
2) List the 'obvious' points to discuss



- Obvious (State the obvious)

- Details (Evidence & Numerical details)
What is the **Total Variation** of the Data?

What is the **Trend** Component

What is the **Seasonal** Component

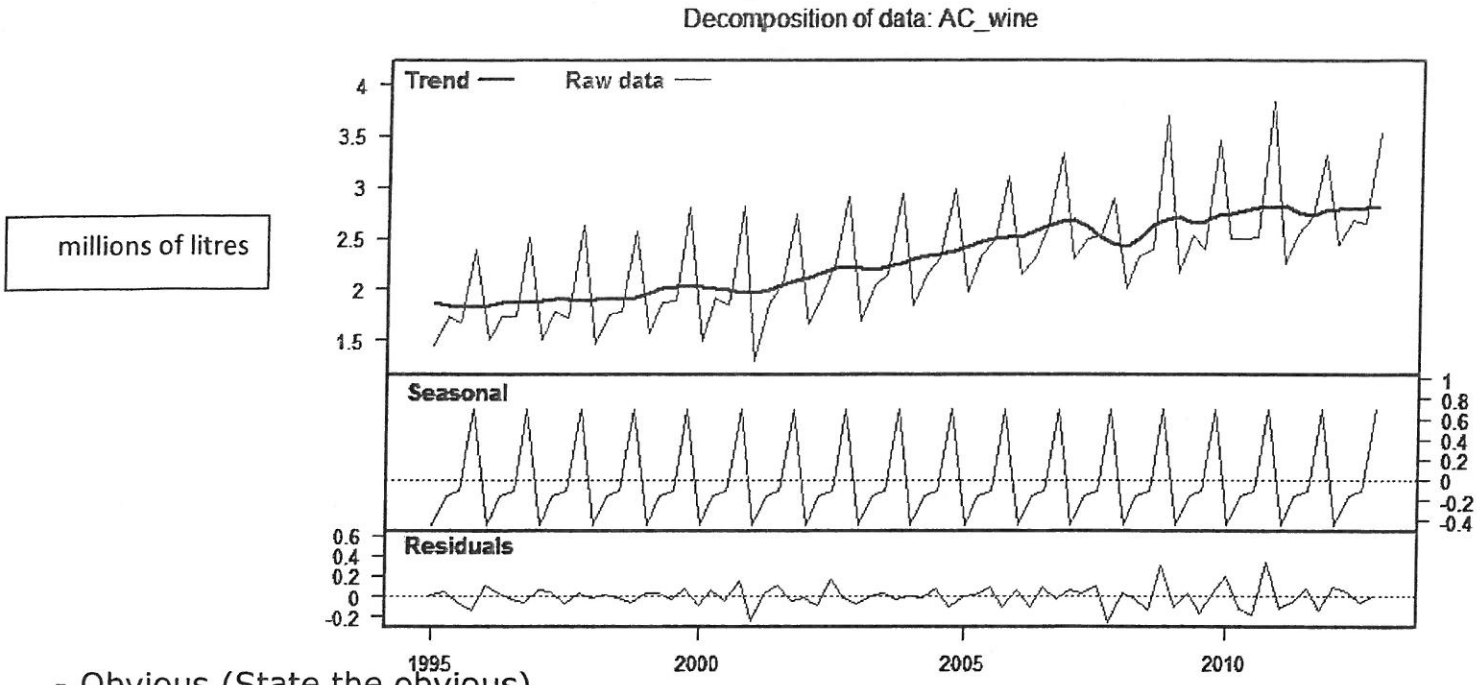
What is the **Residual** Component

How do these compare?

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)



- Obvious (State the obvious)

- Details (Evidence & Numerical details)
What is the **Total Variation** of the Data?

What is the **Trend** Component

What is the **Seasonal** Component

What is the **Residual** Component

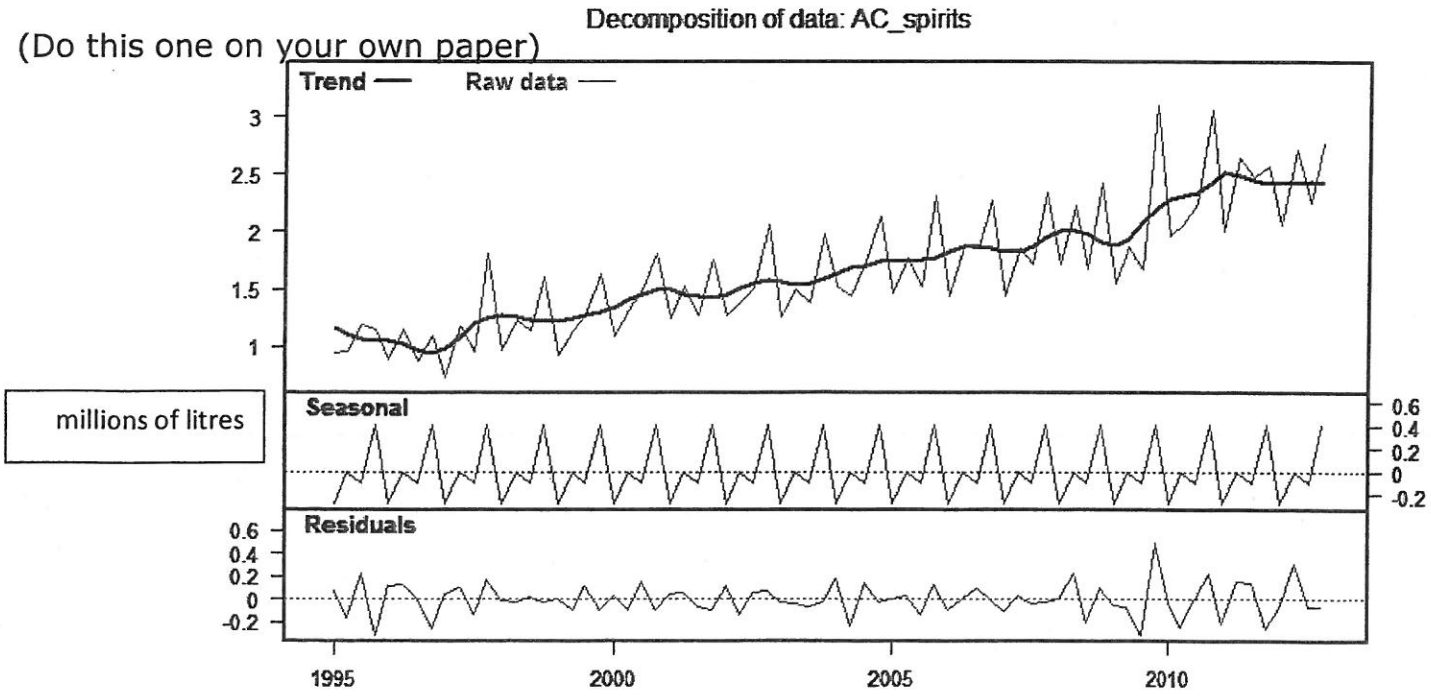
How do these compare?

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

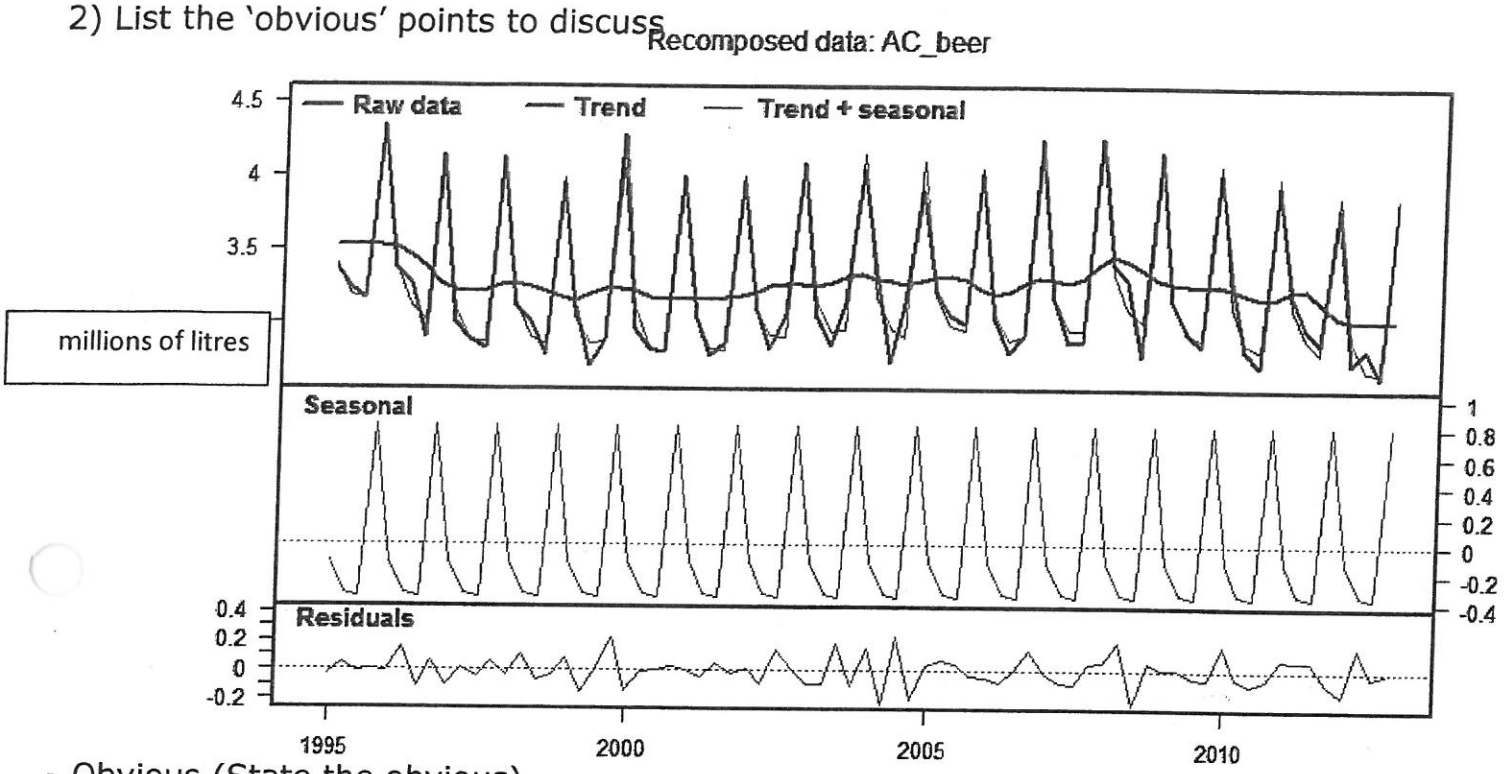
- Relate (To references / research and hypothesis made)

(Do this one on your own paper)



Alcohol consumption in New Zealand by type of alcohol.
RECOMPOSED DATA

1) Graph from iNZight
2) List the 'obvious' points to discuss



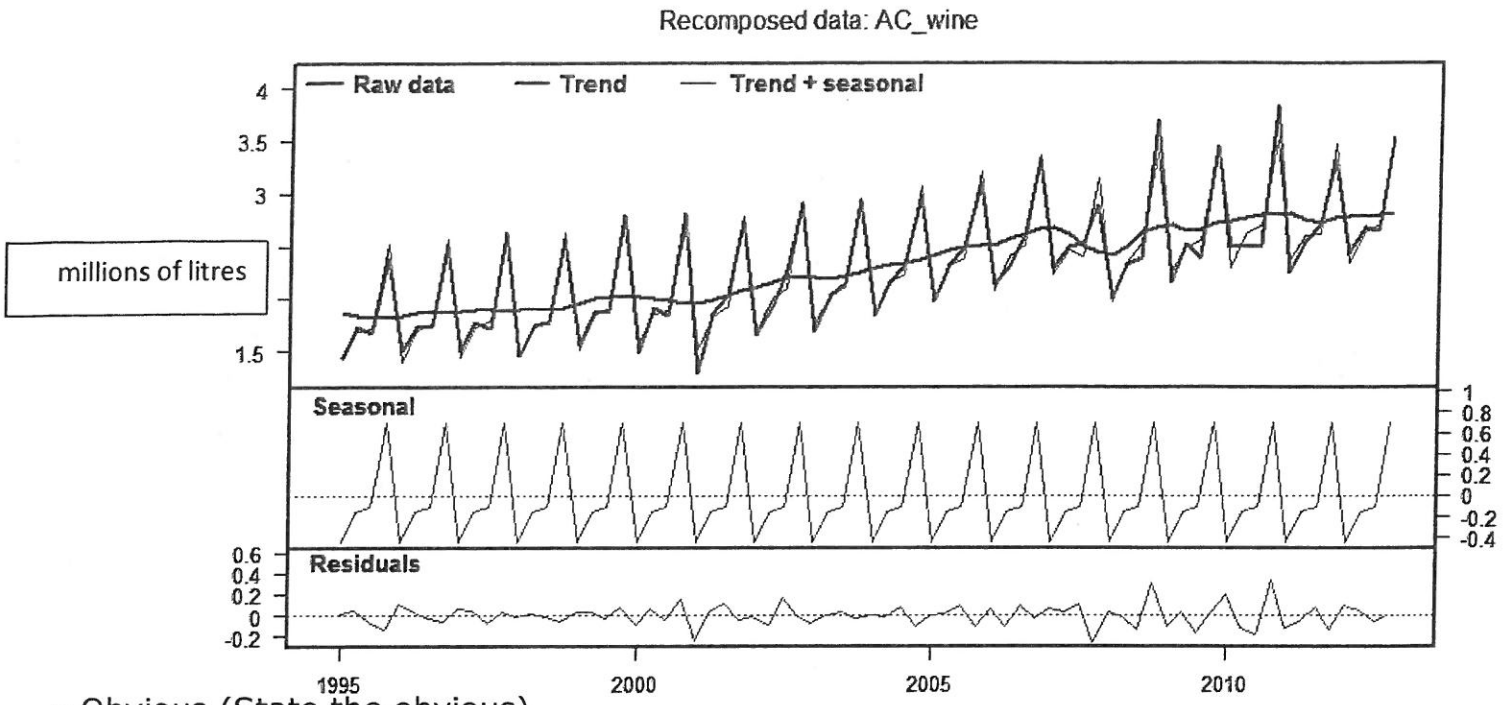
- Obvious (State the obvious)

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)



- Obvious (State the obvious)

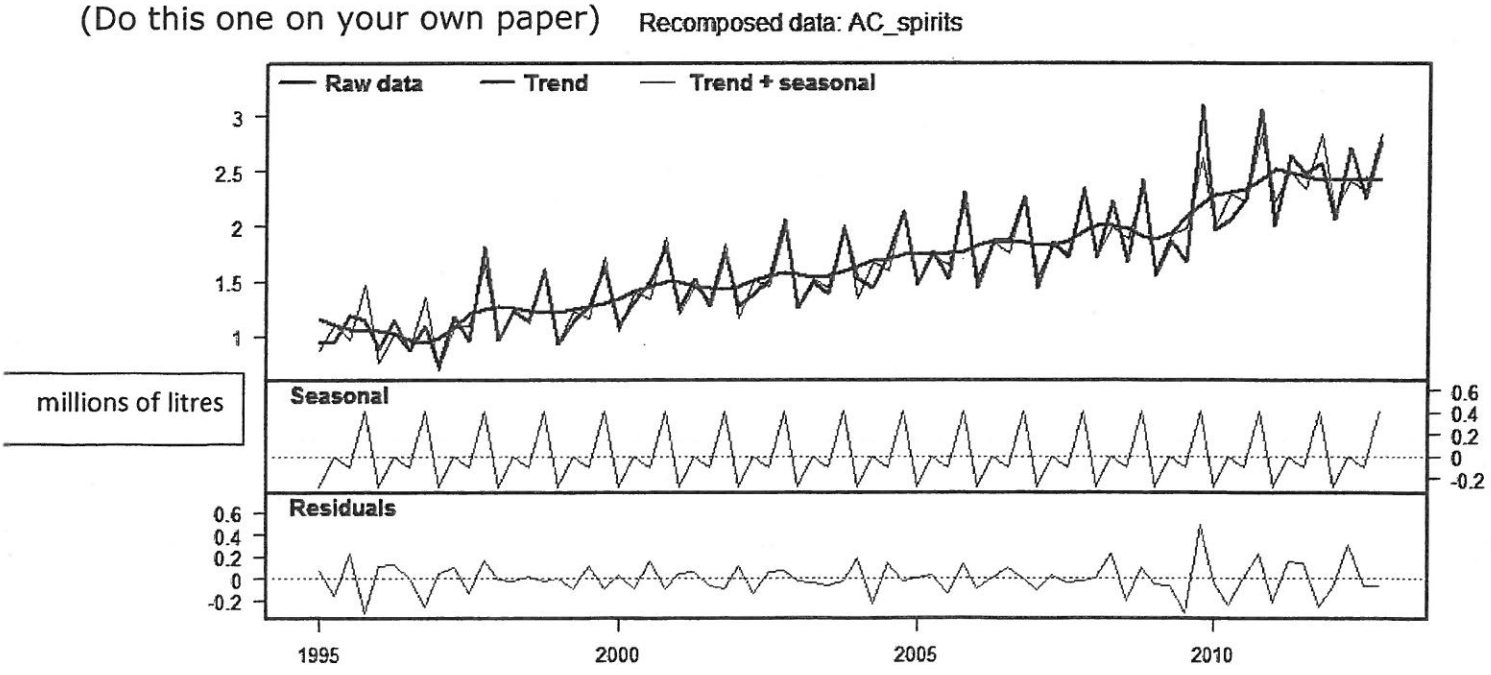
- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

Relate (To references / research and hypothesis made)

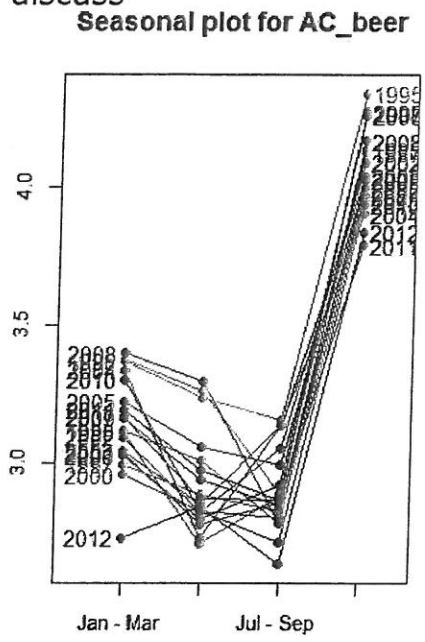
(Do this one on your own paper)

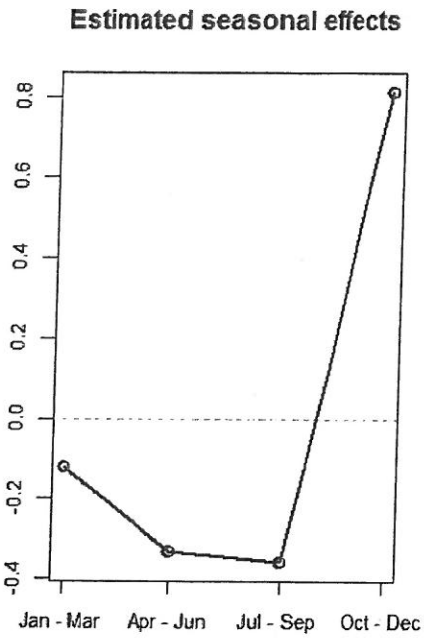


Alcohol consumption in New Zealand by type of alcohol.
SEASONAL EFFECTS

1) Graph from iNZight
2) List the 'obvious' points to discuss

millions of litres





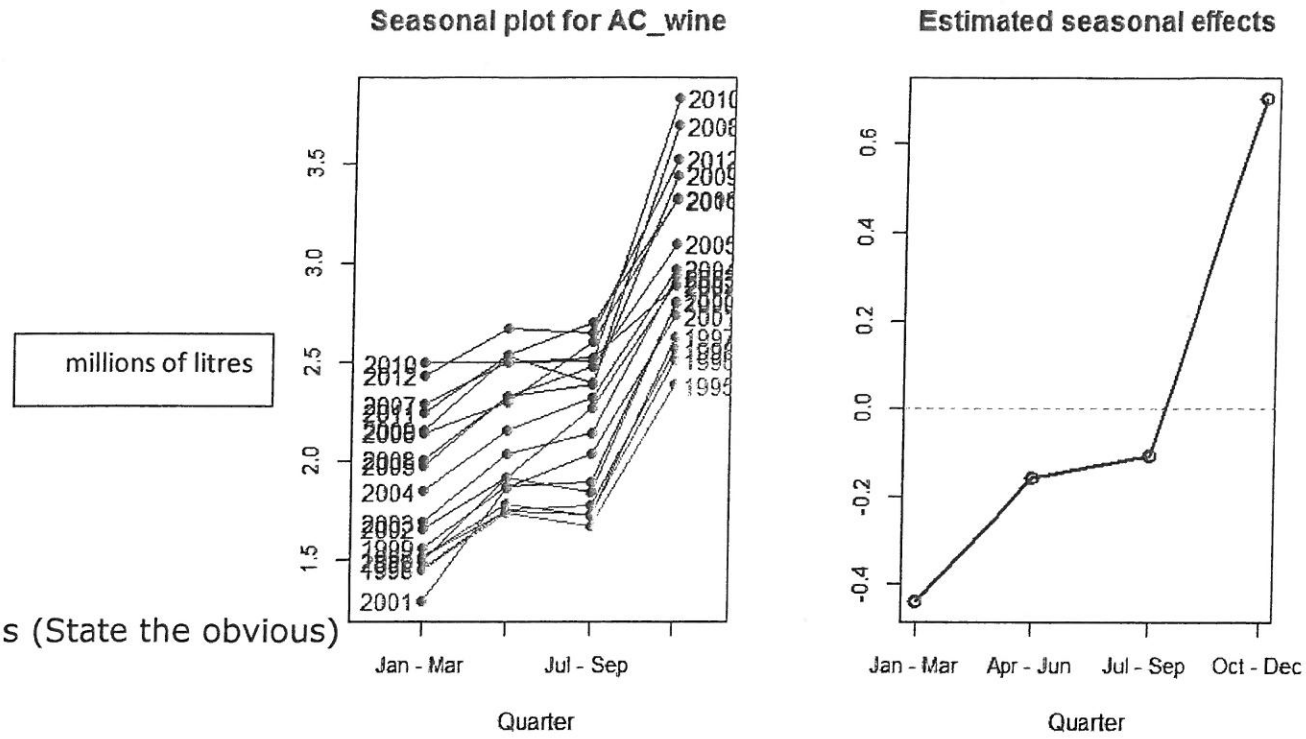
- Obvious (State the obvious)

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)



- Obvious (State the obvious)

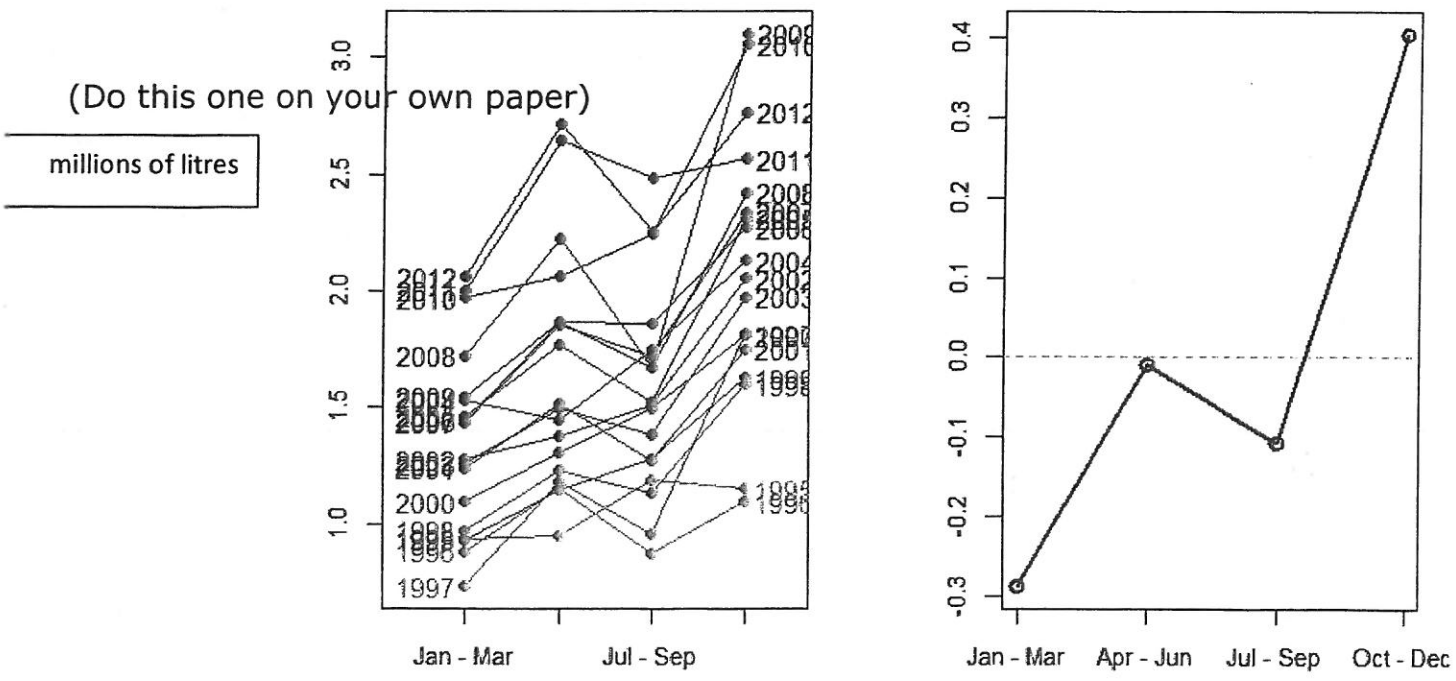
- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

Relate (To references / research and hypothesis made)

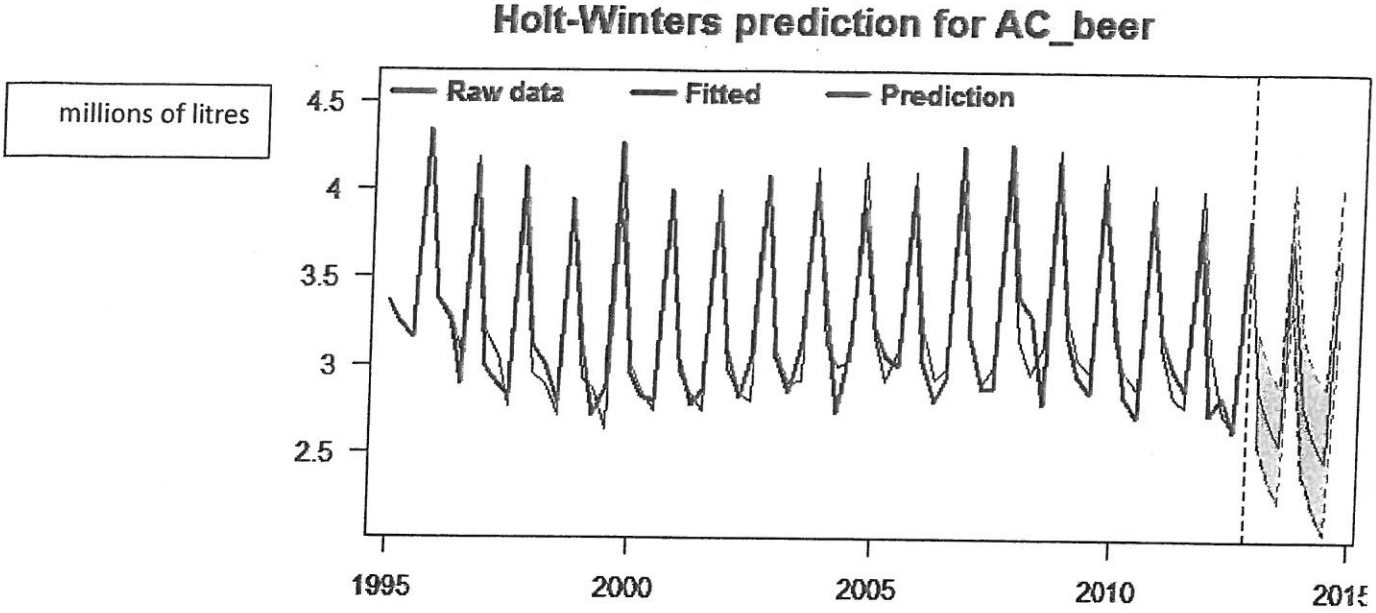
(Do this one on your own paper)



Alcohol consumption in New Zealand by type of alcohol.

PREDICTIONS

1) Graph from iNZight, 2) List the 'obvious' points to discuss



- Obvious (State the obvious)

	fitted	lower 95% bound	upper 95% bound
2013 Q1	2.883499	2.568705	3.198294
2013 Q2	2.685366	2.363712	3.007020
2013 Q3	2.561459	2.230653	2.892264
2013 Q4	3.715067	3.372634	4.057499
2014 Q1	2.792589	2.413301	3.171877
2014 Q2	2.594456	2.199295	2.989617
2014 Q3	2.470548	2.056962	2.884135
2014 Q4	3.624156	3.189627	4.058685

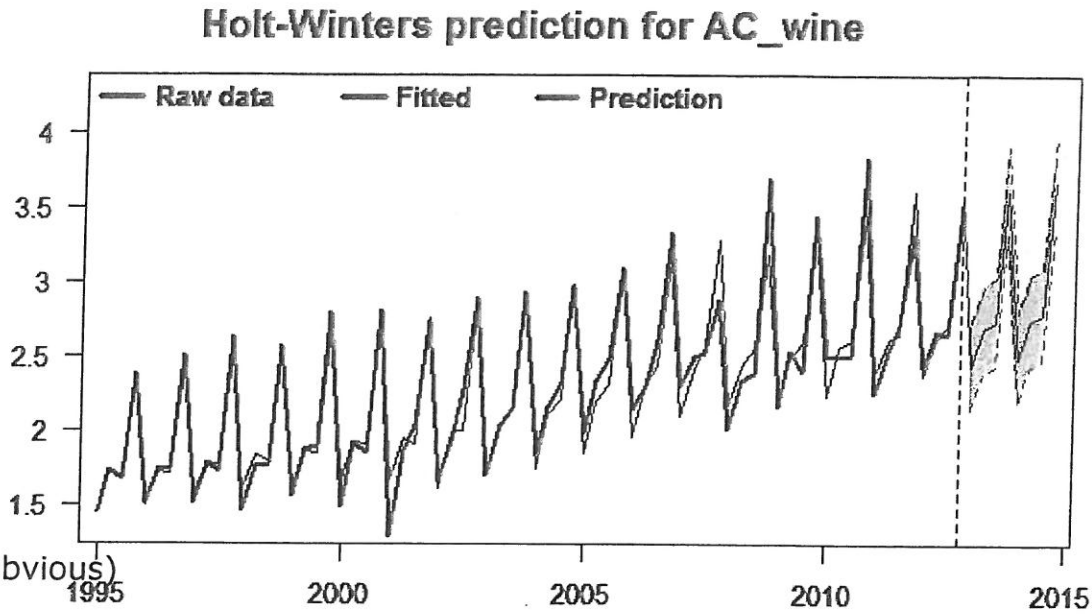
- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)

millions of litres



- Obvious (State the obvious)

	fitted	lower 95% bound	upper 95% bound
2013 Q1	2.435056	2.140961	2.729151
2013 Q2	2.681631	2.386588	2.976673
2013 Q3	2.734007	2.438020	3.029995
2013 Q4	3.614686	3.317757	3.911615
2014 Q1	2.487106	2.177712	2.796500
2014 Q2	2.733681	2.423386	3.043976
2014 Q3	2.786057	2.474864	3.097251
2014 Q4	3.666736	3.354647	3.978825

- Details (Evidence & Numerical details)

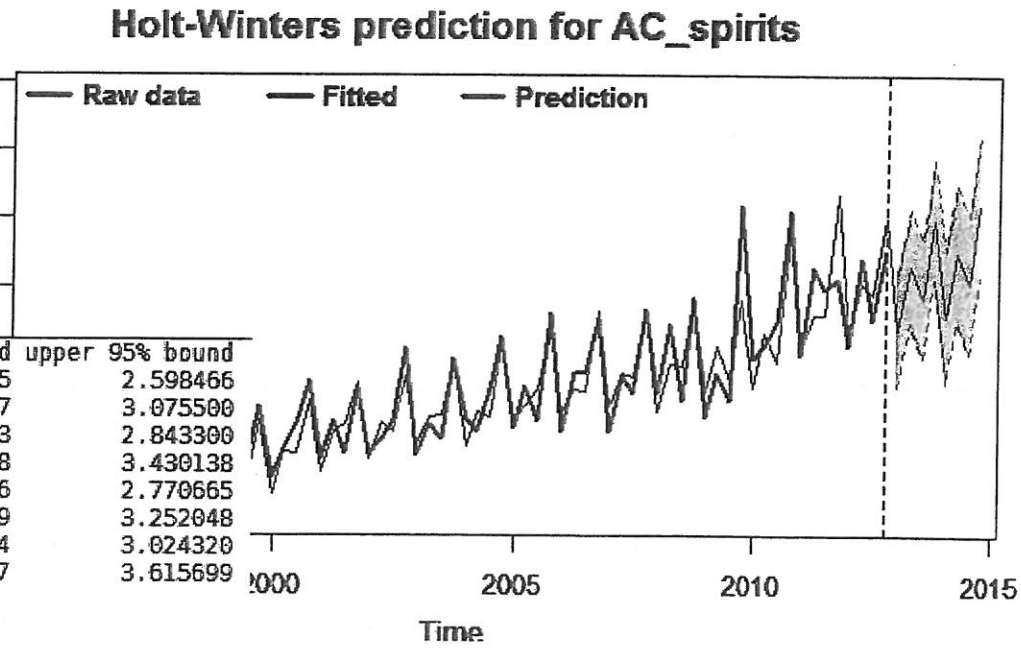
- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)

(Do this one on your own paper)

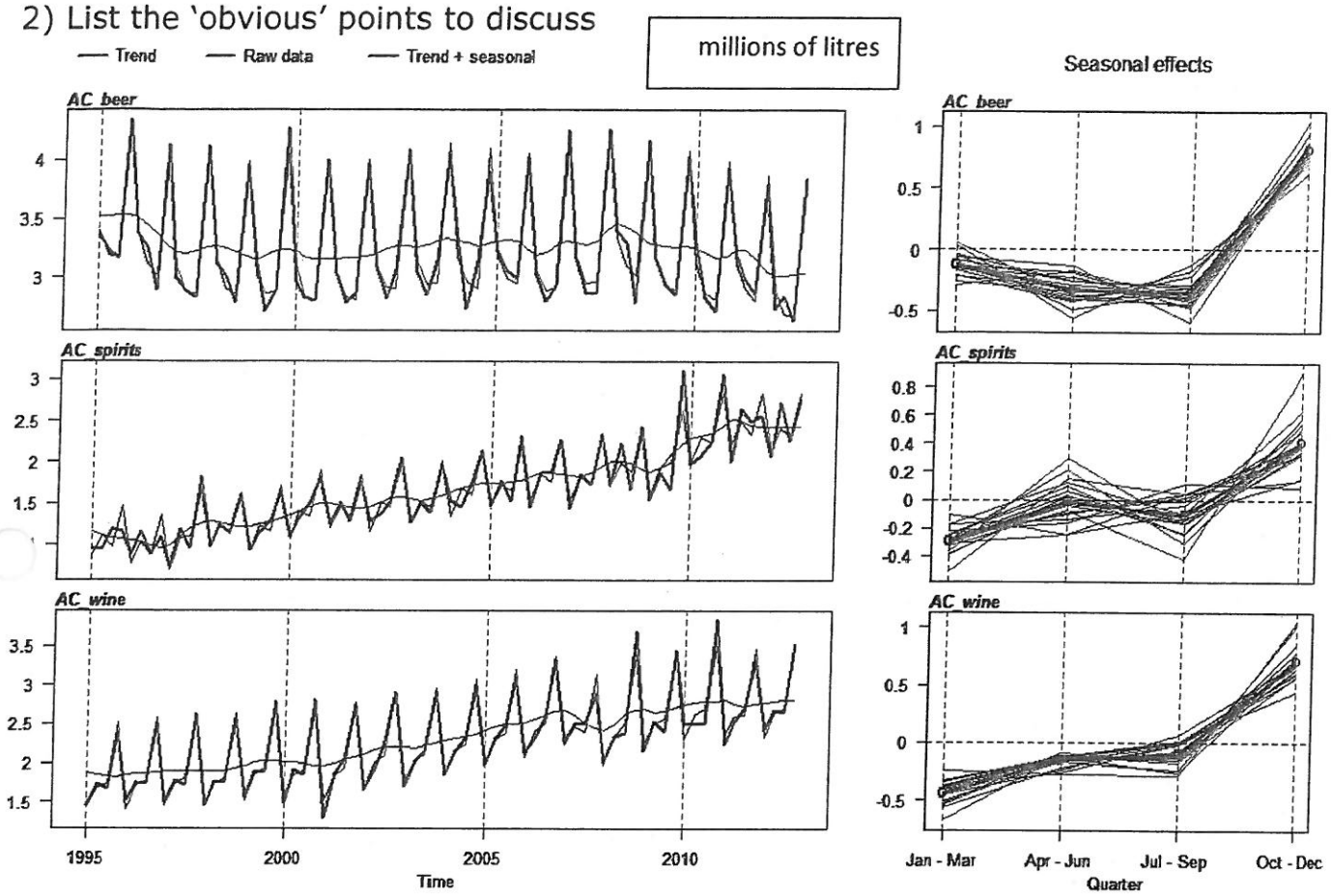
millions of litres



	fitted	lower 95% bound	upper 95% bound
2013 Q1	2.186200	1.773935	2.598466
2013 Q2	2.657483	2.239467	3.075500
2013 Q3	2.418302	1.993303	2.843300
2013 Q4	2.996853	2.563568	3.430138
2014 Q1	2.284961	1.799256	2.770665
2014 Q2	2.756244	2.260439	3.252048
2014 Q3	2.517062	2.009804	3.024320
2014 Q4	3.095613	2.575527	3.615699

Alcohol consumption in New Zealand by type of alcohol.
COMPARING VARIABLES

1) Graph from iNZight
2) List the 'obvious' points to discuss



- Obvious (State the obvious)

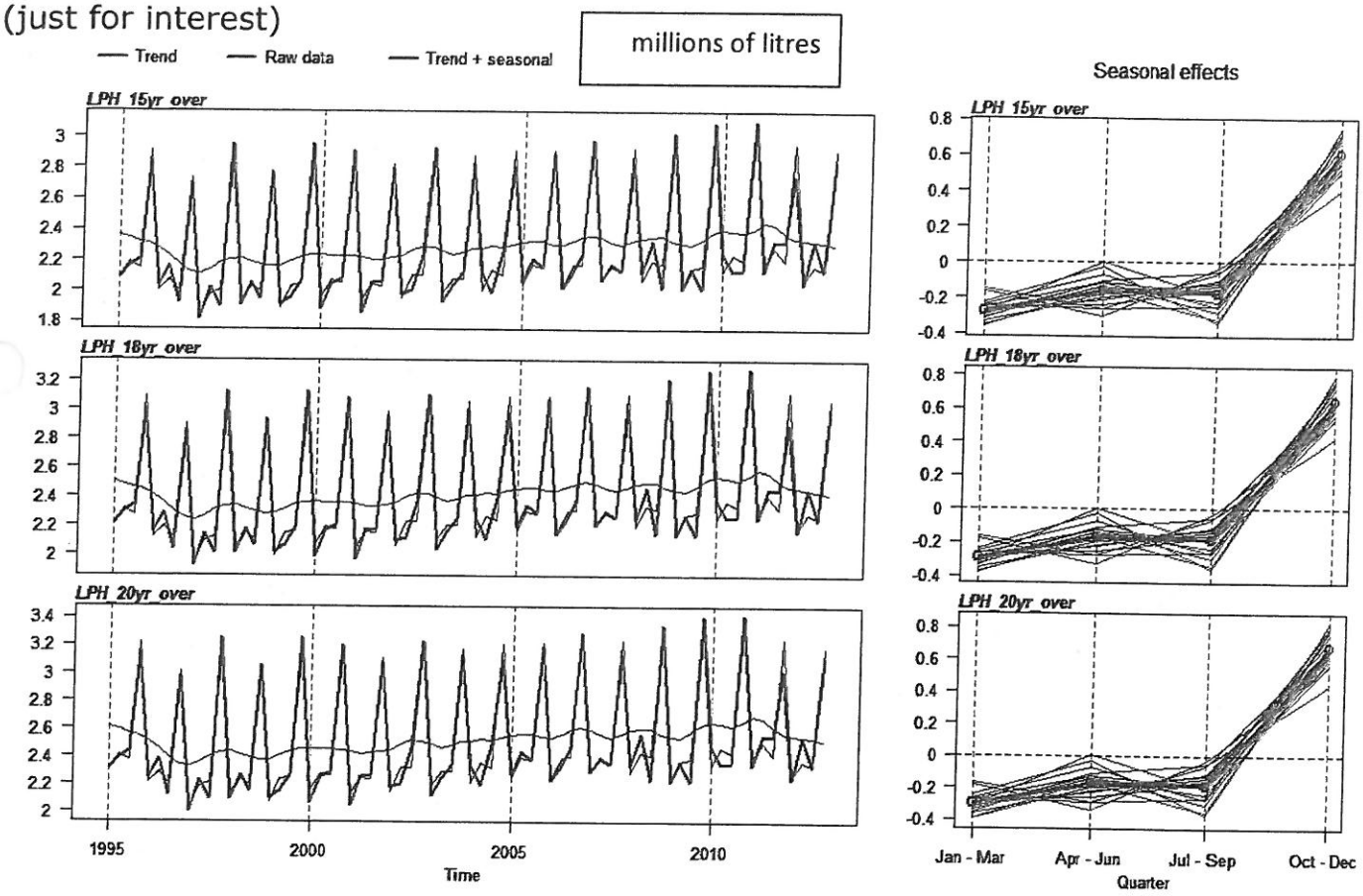
- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)

(just for interest)



- Obvious (State the obvious)

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

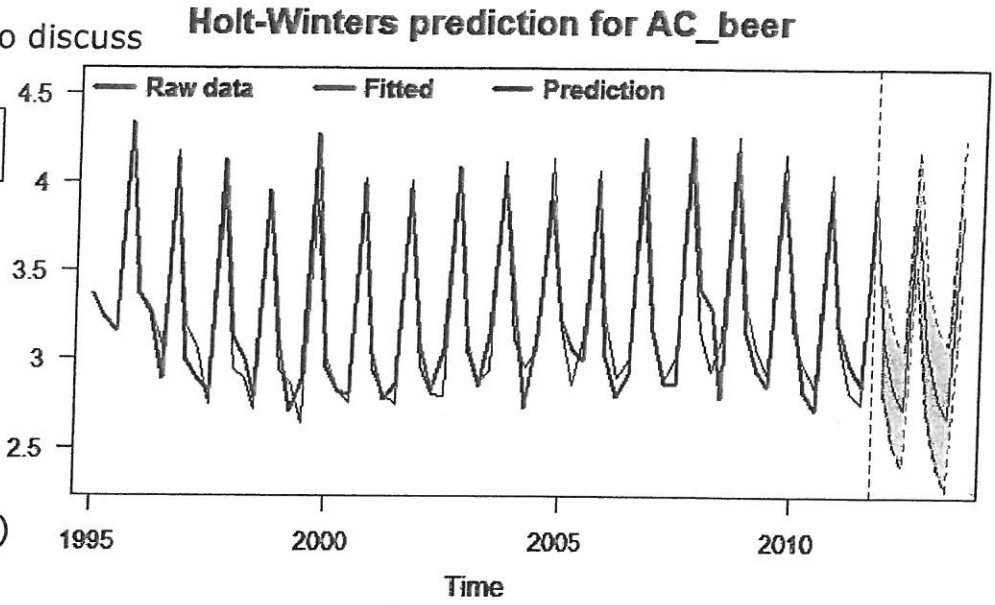
- Relate (To references / research and hypothesis made)

Alcohol consumption in New Zealand by type of alcohol.
TESTING ROBUSTNESS

1) Graph from iNZight
2) List the 'obvious' points to discuss

millions of litres

Quarter	AC_beer	AC_spirits	AC_wine
2012Q1	2.728	2.061	2.433
2012Q2	2.84	2.715	2.673
2012Q3	2.633	2.252	2.647
2012Q4	3.837	2.764	3.526



- Obvious (State the obvious)

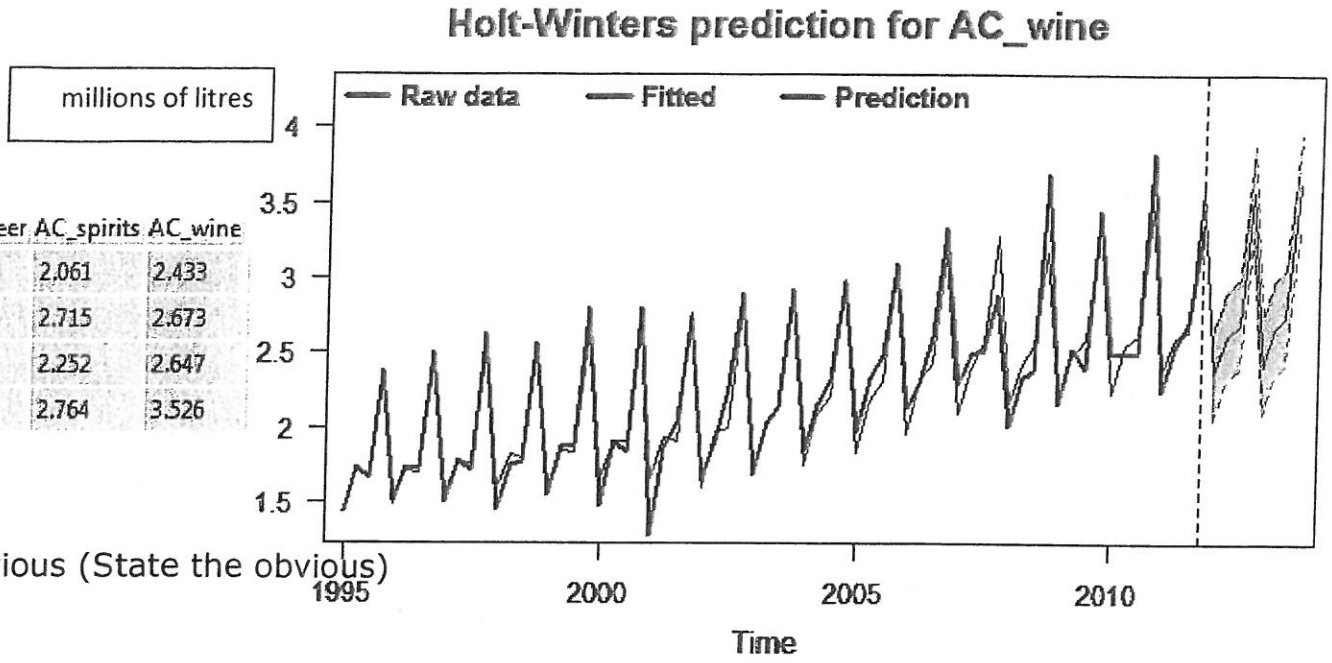
	fitted	lower 95% bound	upper 95% bound
2012 Q1	3.130877	2.822321	3.439433
2012 Q2	2.840001	2.524668	3.155335
2012 Q3	2.724945	2.400910	3.048981
2012 Q4	3.862804	3.528019	4.197589

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)



Quarter	AC_beer	AC_spirits	AC_wine
2012Q1	2.728	2.061	2.433
2012Q2	2.84	2.715	2.673
2012Q3	2.633	2.252	2.647
2012Q4	3.837	2.764	3.526

- Obvious (State the obvious)

- Details (Evidence & Numerical details)

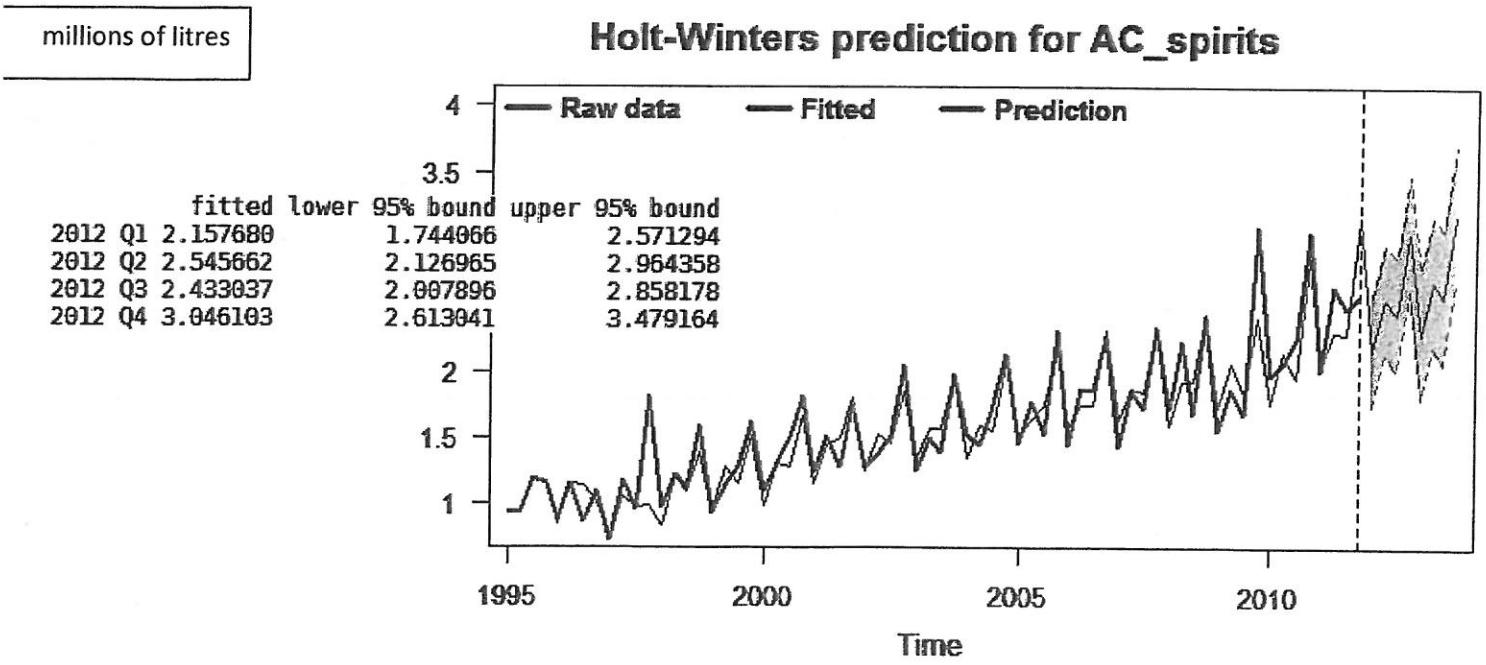
	fitted	lower 95% bound	upper 95% bound
2012 Q1	2.368339	2.066319	2.670358
2012 Q2	2.617801	2.314763	2.920839
2012 Q3	2.692776	2.388723	2.996829
2012 Q4	3.573467	3.268402	3.878531

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)

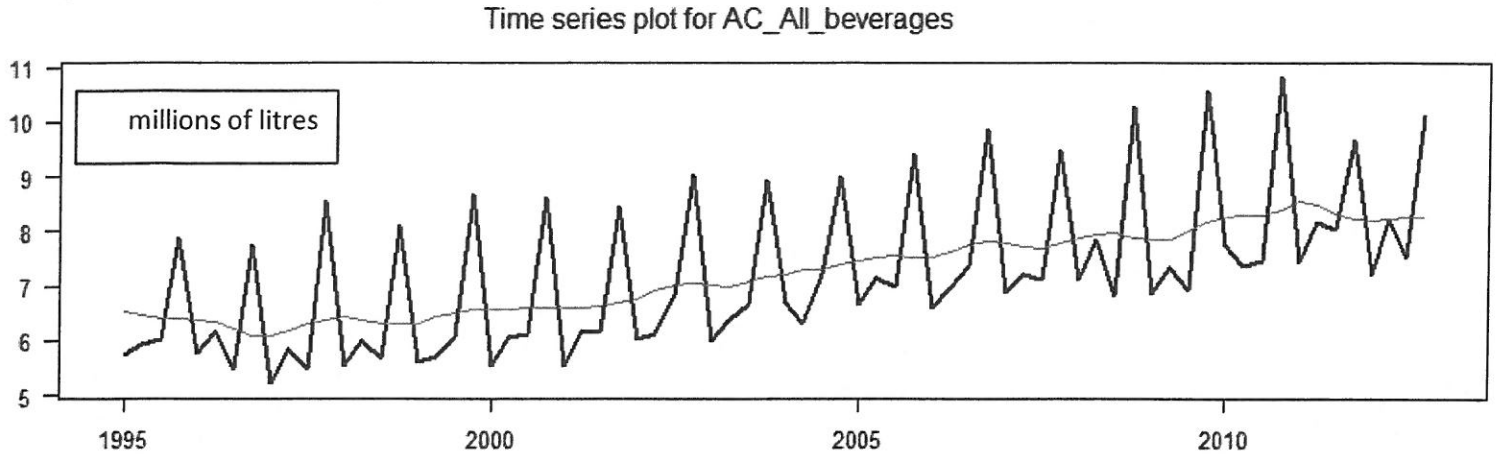
(Do this one on your own paper)



	fitted	lower 95% bound	upper 95% bound
2012 Q1	2.157680	1.744066	2.571294
2012 Q2	2.545662	2.126965	2.964358
2012 Q3	2.433037	2.007896	2.858178
2012 Q4	3.046103	2.613041	3.479164

Alcohol consumption in New Zealand by type of alcohol.
COMBINING VARIABLES

1) Graph from iNZight
2) List the 'obvious' points to discuss



Obvious (State the obvious)

- Details (Evidence & Numerical details)

- Context (Relate to the context. What does this mean?)

- Assumptions (check & discuss any statistical assumptions)

- Relate (To references / research and hypothesis made)

AS 3.8 Investigate Time Series Data AS91580

'Tell the story' of time series data and show evidence of each component of the Statistical Cycle:

(1) Pose a Question
Achieve: Identify a purpose for the investigation.

Merit / Excellence: Research the context and justify the choice of variable(s) to investigate from the data set with respect to a clear purpose for the investigation.

(2) Plan
Using existing data sets & selecting an applicable variable to investigate.

(3) Display
Achieve: Produce plots of raw, smoothed data + seasonal effects. Small labelling errors allowed.

Merit / Excellence: Produce plots of raw, smoothed data + seasonal effects. The appropriateness of the model is justified throughout the entire range of x-values

(4) Analysis
Achieve: Calculate a trend & seasonal effect for one series (calculate in EXCEL or use iNZight) fit a linear model and make predictions. Calculate at least one forecast (not necessarily in context). Correct units given.

Merit: As above but forecast must be in context without errors. Correct units given. It is justified with a discussion on how precise the forecast might be based on the reliability of the trend or the seasonal component. Any comments made must be supported with references to statistical evidence. There is an understanding that the forecasted values are estimates.

Excellence
- **Improvements**: such as alternative models or piecewise models. Fitting an alternative model to specified ranges of x-values could be considered. Any alternative models would have to be justified as being an improvement in terms of how reliable they are in the making of forecasts.

- **Make forecast** (estimates of future data values). Calculate some forecasts in context without errors and with correct units. Comment on the accuracy of forecasts based on the reliability of the trend or the seasonal component (And the influence of trend line, especially variation in the end of the trend line) There is an understanding that the forecasted values are estimates.

-**Robustness:** Test the model & prediction robustness by removing some data the re-predicting. Make a comparison of the actual and predicted values and discuss

- **Comparison of Variables:** Show deeper understanding of model and or context by making a sensible comparison between two related series which add to the depth of the investigation.

- **Combine Variables:** Create another variable to the analysis and investigate the new series (sum, difference or ratio of series). **This should have a purpose and add something to your investigation**. (otherwise don't do it)

- **Compare Excel & iNZight Forecasts:** and discuss difference / similarities and relate to confidence intervals from iNZight

- **Compare Additive & Multiplicative models:** and discuss the applications of the different models and the effects on the forecasts

(5) Conclusion

Achieve: Comment on (describe) trend and seasonal pattern. Describe trend and seasonal pattern not necessarily in context. Quantify trend - (from gradient in EXCEL or Read first and last trend values from graph in INZight)

Merit: Comment on (describe) trend and seasonal pattern in context. Quantify trend - (from gradient in EXCEL or Read first and last trend values from graph in INZight) Comment on accuracy of prediction, fit of the model, consistency of seasonal pattern. INZight can produce confidence Intervals for predictions.

Excellence

- Referencing & Context: Have contextual references throughout the entire investigation to support findings.

- Research: Reflect on the analysis with respect to the background research undertaken and comment on research findings that confirm or dispute analysis

For Excellence: No misunderstanding shown or conflicting statements are present.

Notes:

For the assessment students will be provided with a time series data set containing multiple variables. Background information related to the data set will be provided. Students should be sourcing relevant contextual knowledge about the situation under investigation from places such as the internet, the school or local library, newspapers and magazines. These sources should be referenced in their report.

What to do.....

Make a graph the time series data with trend line. (find the gradient)	**Discussion:** the **Long Term Trend** in very general terms but in context
Decompose the data into trend, seasonal & residual. (work out the % contribution of each component)	**Discuss** relative effect of **seasonal effect vs long term trend** **Residuals** - Any unusual observations - do they warrant further investigation
Recompose the data.	**Discuss** recomposed data and individual data points (above and below average)
Graph the Individual & Estimated Seasonal effects.	**Discuss** the estimated seasonal effects ie what one seasonal cycle is like and possible reason why? Discussion in context **Discuss** the individual seasonal effects and how they may have changed over time.
Make Predictions of the next two cycles of data (with confidence intervals)	**Discuss** the predictions in context (with correct units) **Discuss** the accuracy or margin or error of the predictions
'Compare Series' between different variables that are available.	**Discuss** what you notice about the comparative data series - similarities, differences, possible relationships, reasons, causes, links etc **Discuss** and compare the Trend Lines. **Discuss** and compare the Average Seasonal Effects (red line) and the Seasonal Effects for each cycle (gray lines)
Combine variables to make a new variable into the series and analyse.	**Discuss** why the new variable has been added. **Discuss** the further insight and information provided by the new variable
Compare the 'additive and the 'multiplicative' models.	**Discuss** & compare & contrast the two models. Which may be more appropriate and why? Compare forecasts.

AS 3.8 Time Series Checklist

Introduction
- ☐ Description & Investigative Question "I will investigate..."
- ☐ Source of Data given
- ☐ Aim / Interest for investigation
- ☐ Variables defined & described
- ☐ Hypothesis 'I think that... because...
- ☐ Research findings summarized
- ☐ Research the context and justify the choice of variable(s)

Trend
- ☐ Graph of raw and smoothed data
- ☐ Vertical axis labels added with units
- ☐ Describe the trend **in context: (with numerical values)**
- ☐ Discuss further aspects of interest

Decomposed data
- ☐ Graph of decomposed data
- ☐ Calculate and discuss the relative size of each of the components (trend, seasonal & residual)
- ☐ Discuss further aspects of interest

Recomposed data
- ☐ Graph of recomposed data
- ☐ Comparing the Raw Data with the 'Recomposed Data' and discuss.
- ☐ Discuss further aspects of interest

Seasonal effects.
- ☐ Graphs of individual & average seasonal effects.
- ☐ Vertical axis labels added with units
- ☐ Describe & discuss the Individual seasonal cycles **in context (with numerical values)**
- ☐ Describe the Estimated Seasonal cycle **in context (with numerical values)**
- ☐ Discuss further aspects of interest

Predictions
- ☐ Table of predictions
- ☐ **At least one actual forecast with correct units & rounding (in context)**
- ☐ Discussion in context and predictions rounded sensibly.
- ☐ Discuss what the prediction error means - accuracy of forecasts (variation)

Conclusion
- ☐ Answer the Statistical Question, with references & hypothesis,
- ☐ Evaluation and summary

Excellence (some ideas... not all are needed – Quality, not Quantity)
- Detailed thoughtful discussion:
- Explanation for variations in the trend line,
- Non-linear trend models, Piecewise models, Comparison of iNZight & Excel'
- Recent variation in iNZight trend line...
- How might the forecasts be used (and who might use them) - discuss.
- Has the end of the trend line has been influenced by the position in the seasonal cycle of the end point?
- Testing the robustness of the model by removing recent data, re-predicting
- Investigate links between variables by comparing them. Thoughtful choice of variables to compare is needed to be discussed
- Discuss what you notice about the comparative data series - similarities, differences, possible relationships, reasons, causes, links etc
- Combine variables together to create a new variable to investigate, **think carefully as to the PURPOSE of the new variable.** What are you wanting to show? Why is the new variable useful?
- Discuss why the new variable has been added.
- Discuss the further insight and information provided by the new variable