Dover Ferries

An annotated exemplar for Time Series

Introduction

This report relates to the travel by sea to Europe of people from Britain.

I have chosen to concentrate on ferry travel from the port of Dover, which is part of the Thames and Kent area in the data. Dover is the closest point in England to Europe and has historically been the main point of travel across, helped by good rail links from London.

Ferries from Dover go to Calais and Dunkirk in France (see map above). They also used to go to Boulogne (near Calais) but this stopped in 2010. The Ramsgate to Ostend route near to the Dover routes closed in 2013, which seems to have affected the Dover numbers as people who used to go to Ramsgate now often use Dover instead.

The Dover–Calais route is very old and busy. At the moment it includes 15 catamaran crossings each day, and 20 to 30 trips by ordinary boats. Many people take cars across on the ferries. Passengers include those going to other places in Europe or just to the Calais area (sometimes even just for the day). Travellers include a lot of tourists and holidaymakers but also a freight component with a large number of trucks.

Dover to Dunkirk has been going since 2000, and has the same mix of tourists and freight, but with fewer sailings.

I shall see if we can usefully predict the amount of ferry passengers from Dover for the next years.

Interested Parties

People interested in the results of this analysis might include:

- Tourist businesses in the Dover area as they depend on the traffic from the ferries (the truck drivers tend to just drive through).
- Dover District Council and the Port of Dover, as they derive revenue from the ferries.
- Rail companies, as a lot of the travellers arrive by rail.

Predictions of the amount of ferry traffic in the future is therefore of some importance to them. I will use my analysis to make such a forecast that a local business might be able to use.

A reason why someone might use the analysis is required. Bring it up again at your conclusion.

Events affecting Dover Ferries:

1995: Channel tunnel opens.

1998: Channel tunnel reaches full operation, which takes people away from sea routes.

2010 2nd Quarter: the Icelandic volcano Eyjafjallajökull erupted, and more people went by sea as air travel was disrupted.

2010 3rd Quarter: Dover to Boulogne route stopped.

2013 2nd Quarter: Ramsgate to Ostend route stopped.

> This is a time series analysis. Dates (times) are important.

Sources used include:

Information on Dover: https://en.wikipedia.org/wiki/Dover

Dover District Council: https://www.dover.gov.uk/Home.aspx

Port of Dover: https://www.portofdover.com/

Ferry information: http://www.eurodrive.co.uk/

Channel Tunnel: https://en.wikipedia.org/wiki/Channel_Tunnel

Passenger reports from the UK government's statistics department: https://www.gov.uk/government/statistics/final-sea-passenger-statistics-2014

English school holidays: https://www.schoolholidayseurope.eu/england.html

> Every important source used should be cited. Do **not** make a massive list – just because some page refers to your context a little bit does not mean it is worth listing.
>
> You do not need to use formal citations and footnotes like an academic paper. That can wait to university. Just list the source (full book name, link to web page or citation of a personal interview) and a brief note of what it adds to contextual knowledge.

Data used

The data was from https://www.gov.uk/government/statistical-data-sets/spas01-uk-international-sea-passengers set SPAS0105.

> Do not print out the data itself – the purpose of your analysis is to prevent someone having to read it all.

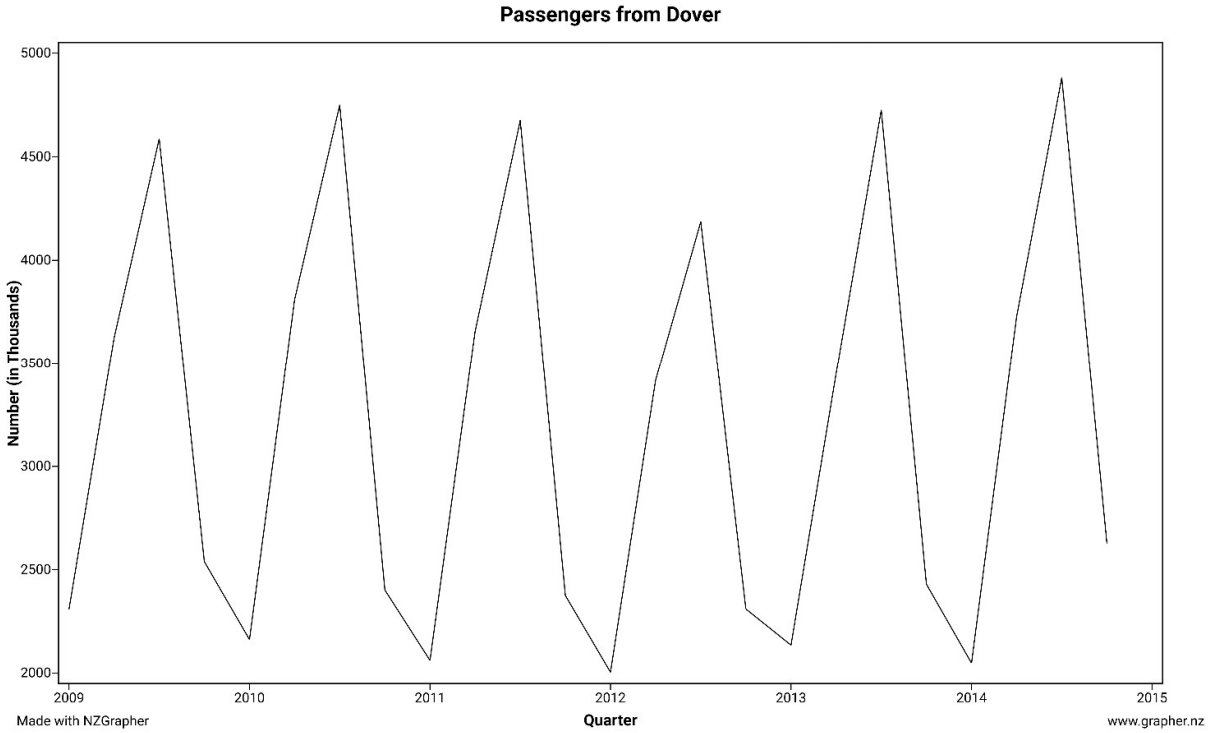
The data set gives the number of passengers in thousands. I have left those figures as they are for the graphs, but used the actual numbers for my report.

> It is **very** common for students to miss issues like this. Always check thoroughly that you know exactly what is being recorded in your data set. This is an automatic fail if you get it wrong.

Analysis

While the report doesn't need to be a book length, it is considered good form to break reports into clearly separate sections (introduction, each separate analysis, conclusion) to help readers keep their place.

The graph below shows my primary time series – all ferry passengers from Dover from the start of 2009 to the end of 2014, by quarter.



You might think the time series doesn't do very much, but this is not a problem. Do not search to find the most "interesting" series to get M or E – odd things happening are just as likely to trip you up as give you better comments. Provided you can find some features to discuss, plain graphs are fine.

Looking at the graph there does not appear to be any obvious long term trend.

There is a clear seasonal pattern, with numbers spiking in the third quarter of each year.

The peaks and troughs vary a bit, but the peak for the third quarter in 2012 is clearly much lower than the other peaks.

Lead in with what you see – trend, seasonal pattern and then variations in those – and only then move to numbers.

I added a long term trend to remove the seasonal effects, shown on the next page.

The long term trend starts at 3,270,000 per quarter and ends at 3,420,000. Over the 16 quarters that averages as (3,420,000 – 3,270,000) ÷ 16 = 9,375 extra people each quarter. In other words 37,500 more each year. In the context of 12,000,000 passengers each year it is a very small percentage change (0.3% in fact).

Always put numbers into the context.

Show your working.

Your gradient use the **long term trend values**. Make sure you get whether it is positive or negative correct.

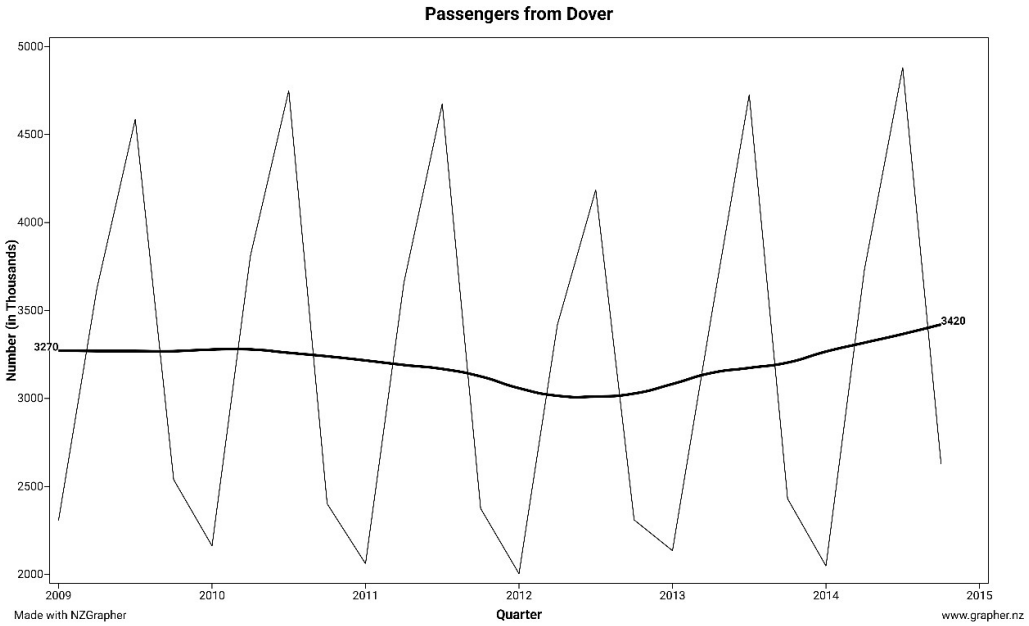
Be very careful to make sure you use what units your data is in.

There is one low year of passengers in 2012. This makes the second half of the trend appear to be rising steadily, whereas it is really just returning to the very long term trend value around 3.3 million passengers.

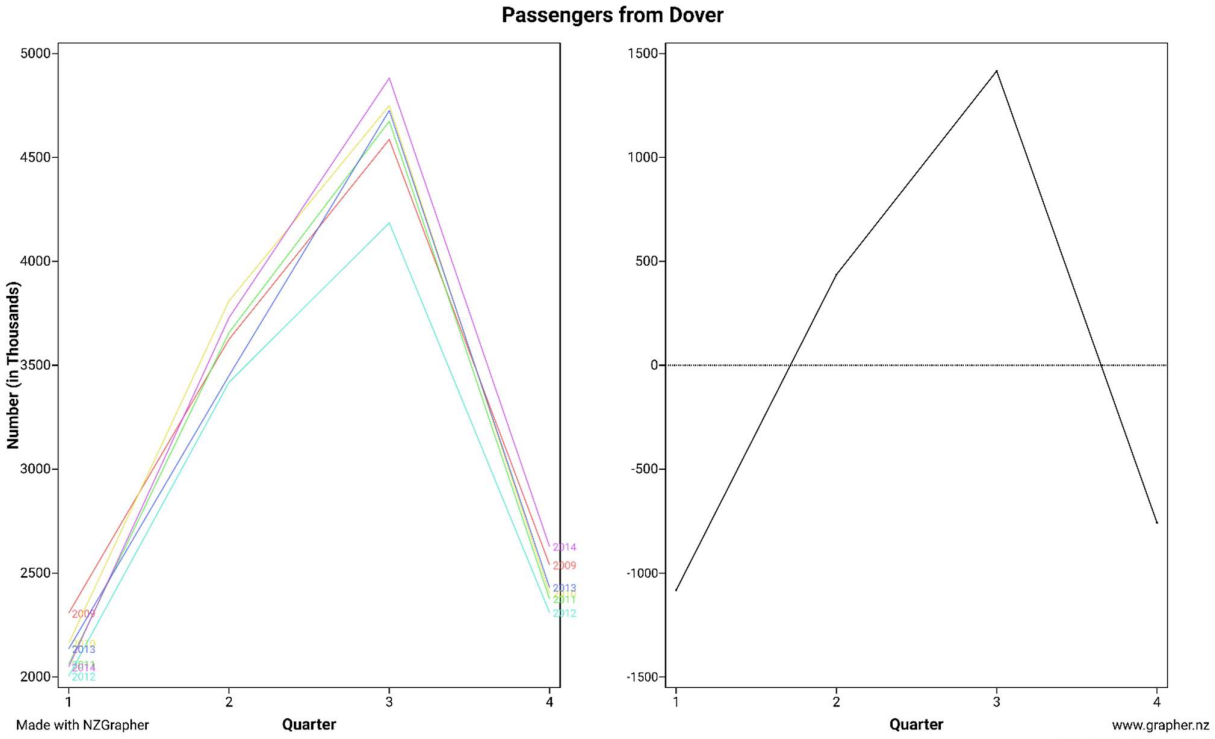
> Excellence level thinking : a "rising" trend might just be a return to normal. Or it could actually be a rise.

My research shows that the 10 year trend in ferry passengers from England is slowly downwards in general as more people fly for their holidays. My graph shows such as small increase (0.3%) that it is not incompatible with a longer term downwards trend.

> Is it what we see what we expect from our research of the topic?



Seasonal Pattern



There is a very strong and consistent seasonal pattern to the data, with each year showing a low in Quarter 1 (just over a million passengers less than trend), a much higher result in Quarter 2, a peak in Quarter 3 (nearly 1.5 million above trend) and then a Quarter 4 value that is almost, but not quite as low as Quarter 1.

> Lead in with what you see, and only then move to numbers. Make sure your numbers are in terms of the **actual numbers**, not the ones in the data set (here in thousands).

These show that on average there is a difference of 2,500,000 passengers between the Q1 trough and Q3 peak. The summer peak is over twice the value of the winter trough.

> Always put the **meaning** of your numbers into context.

This pattern is what you would expect for ferry routes that have a lot of tourists. There is always a flow of people on business and even in winter a few tourists, but there is a huge spike in tourists for the summer months.

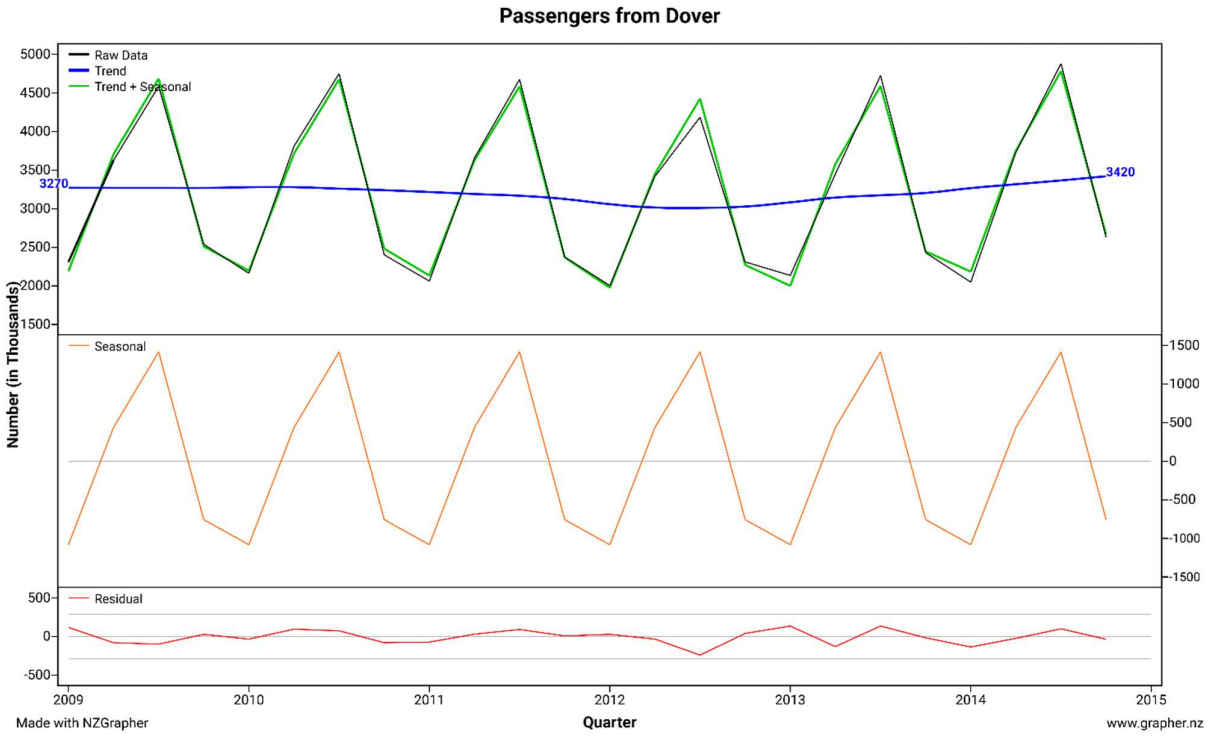
School holidays in England are July and August mostly, that is Quarter 3, which could explain why Quarter 3 is consistently the peak. There is also a school break in May and people who want to travel in good weather avoiding the rush will largely travel in Quarter 2.

> A simple explanation of why the seasonal pattern occurs shows you putting the result in context.

Only one value is significantly different to the usual seasonal pattern, which is the unusually low 2012 peak. As I can't read French it was hard to find out why, but it could be as there were financial issues with a French ferry company which took at least one ferry out of operation that year (https://en.wikipedia.org/wiki/MS_C%C3%B4te_des_Flandres).

> What you find when you poke a little bit more!

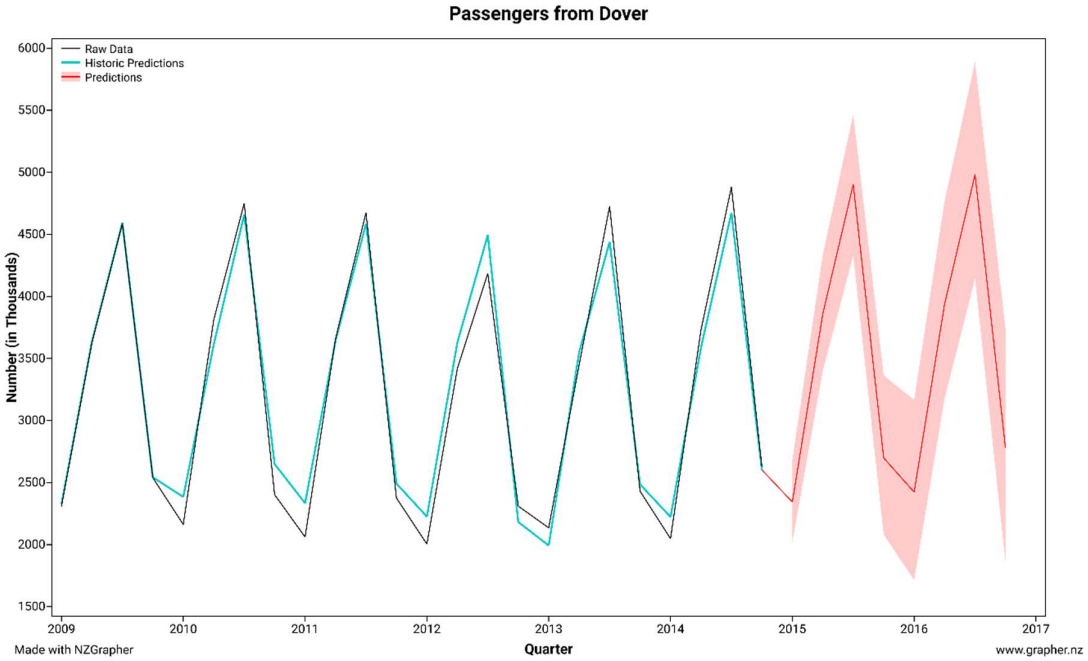
Quality of the Model



As we can see, the model fits the data extremely well, although better in the first half than the second. The residuals are within a band of ±5,000 passengers, which is ±0.2% difference.

> Generally the models don't fit this well, so this section would normally be a lot longer as the deviations of the model from the data are discussed.

Predictions



The model extends into the future with a very small error, which is not surprising given how little the model deviates from the data.

The table:

Time	Min	Prediction	Max
2015Q1	2034.7	2343.3	2674.3
2015Q2	3409.8	3859.7	4315.7
2015Q3	4345.1	4903	5486.6
2015Q4	2055	2698.3	3352.1
2016Q1	1721.2	2422.8	3152
2016Q2	3177	3939.2	4736.9
2016Q3	4097.3	4982.5	5900.4
2016Q4	1798.4	2777.8	3733.2

So the prediction is that the third quarter peak for 2015 is from 4.35 million passengers to 4.32 million passengers, with 3.86 million the number the model predicts.

You must make predictions – it is a compulsory part of the standard.

Round your numbers sensibly, because five significant figures is silly in this context.

There are errors associated with these predictions, as with any extrapolations, and they will tend to increase the further one goes into the future.

Yes it is obvious. Say it anyway.

I suspect this prediction is a tiny bit on the high side because:

- The Holt-Winters model adds the typical seasonal effect onto the long term trend, but emphasises the more recent trend over the less recent. In this case the more recent trend is artificially upwards due to the unusually low 2012 result and may disappear once numbers get back to the 2011 values.
- My research shows that the long term trend is expected to be downwards.

Discussion of the accuracy of the prediction is required at Merit. Ideally that should include both the mathematical aspect of how the trend is calculated and your research.

Otherwise my research shows little reason to think there are major changes in the near future for the ferries. The Channel Tunnel and cheaper flights, which are likely a main reason for the slow downward trend, are not changing. No new ferries are likely to open.

There is always the chance that something big will happen (terrorism affecting people's desire to fly or some action will stop them wanting to visit France etc). That is much more likely to affect the summer quarters. In the past even such large events, such as the eruption of the volcano Eyjafjallajökull, have tended not to have much long term consequence.

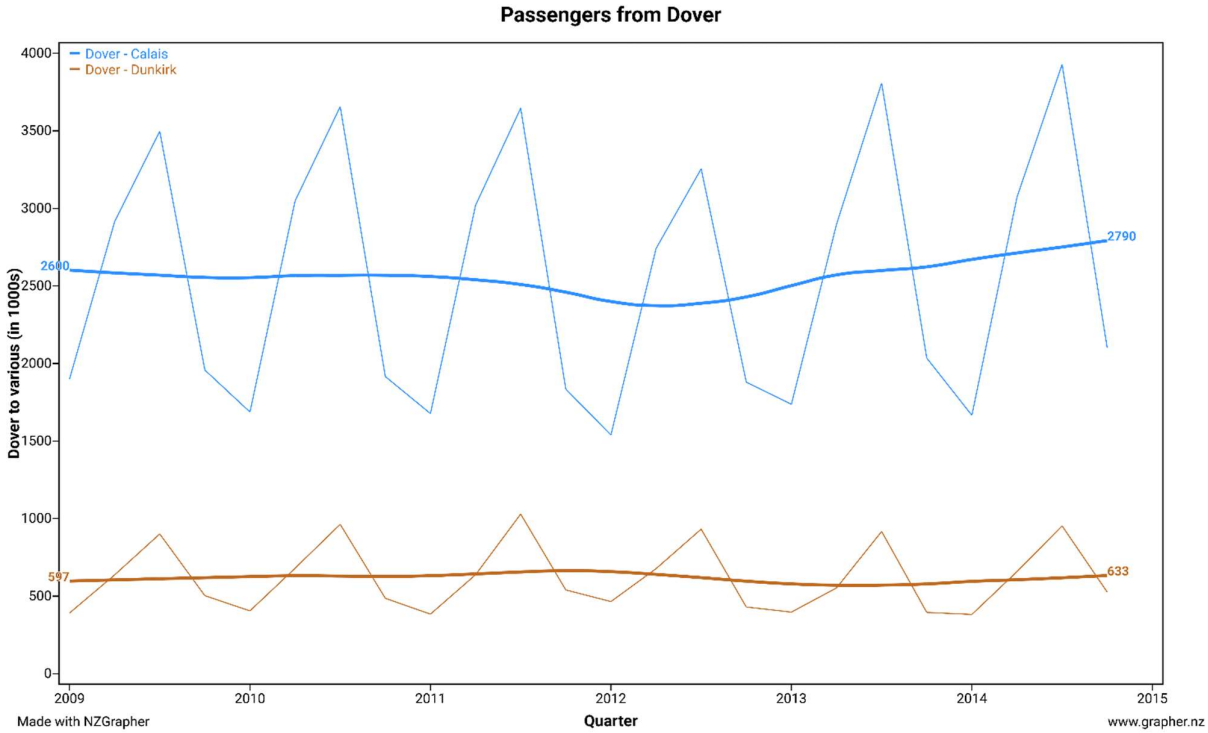
> Discuss the likely factors that might affect your predictions, but keep it brief. A few examples of the major issues is sufficient.

All things considered, I feel that businesses could rely on the prediction within the error margins given.

Further Analysis

> Once you finish the compulsory parts you can look at doing extra analysis. Doing this might get you a higher grade, but only if the main part is strong. You cannot cover for a poor report by doing something extra.

While businesses in Dover don't care which route passengers take, there might be some hidden information underneath, so I separated out the travel by the destination ports of Calais and Dunkirk (Travel to Boulogne stopped in late 2010 and was never very important anyway):



We can see that trend for Dunkirk is very flat indeed, and most of the increase is due to those heading to Calais.

Whatever affected the summer of 2012 clearly was not on the Dunkirk route, which stayed normal.

> While you might do a quick numerical analysis, there is little point showing all the figures like you do with the main analysis.
>
> Don't worry if it doesn't show much interesting. The thing is to show that you can identify possible areas on interest.

Conclusion

You need a conclusion, which you use to repeat your main points clearly. Don't start adding any new material here – if you think of something else for your analysis, go back and add it there.

The number of passengers using the Dover ferries shows a very strong seasonal pattern, underneath of which is a tiny upward trend.

The spring and summer quarters have twice as high passenger numbers as the autumn and winter quarters, which is largely the effect of the much higher number of tourists in the good weather.

The general trend and seasonal pattern, in context.

The seasonal pattern and trend are so strong and regular that businesses in the area could make fairly strong predictions based on them, assuming no major changes in the economy or travel routes.

The overall report here is eight pages, but that is made longer by the annotations, the generous use of white space in the layout and by discussing every part thoroughly.

This is more than long enough: precise explanations with clear graphs are more important than length.

Other ideas that could have been explored at Excellence include:

- a (quick) comparison with a similar port;
- seeing if there is a link between some other economic criterion and the travel numbers (GDP, population, travel numbers, tourism numbers).

Remember further analysis is about looking at some possibly extra interesting feature or some explanation for the main analysis, this is not about repeating the whole process. Do not do more than a couple of such extra investigations.