

Dover Ferries

An annotated exemplar for Time Series

by Man vs Maths

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.

A lot of businesses in the Dover area depend on the traffic from the ferries, mostly tourists (as the truck drivers tend to just drive through).

Explanation to put the data into a context. It should **not** be an essay. While it is good to do this at the front that does not mean that the later analysis should ignore the context.

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>

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 huge 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

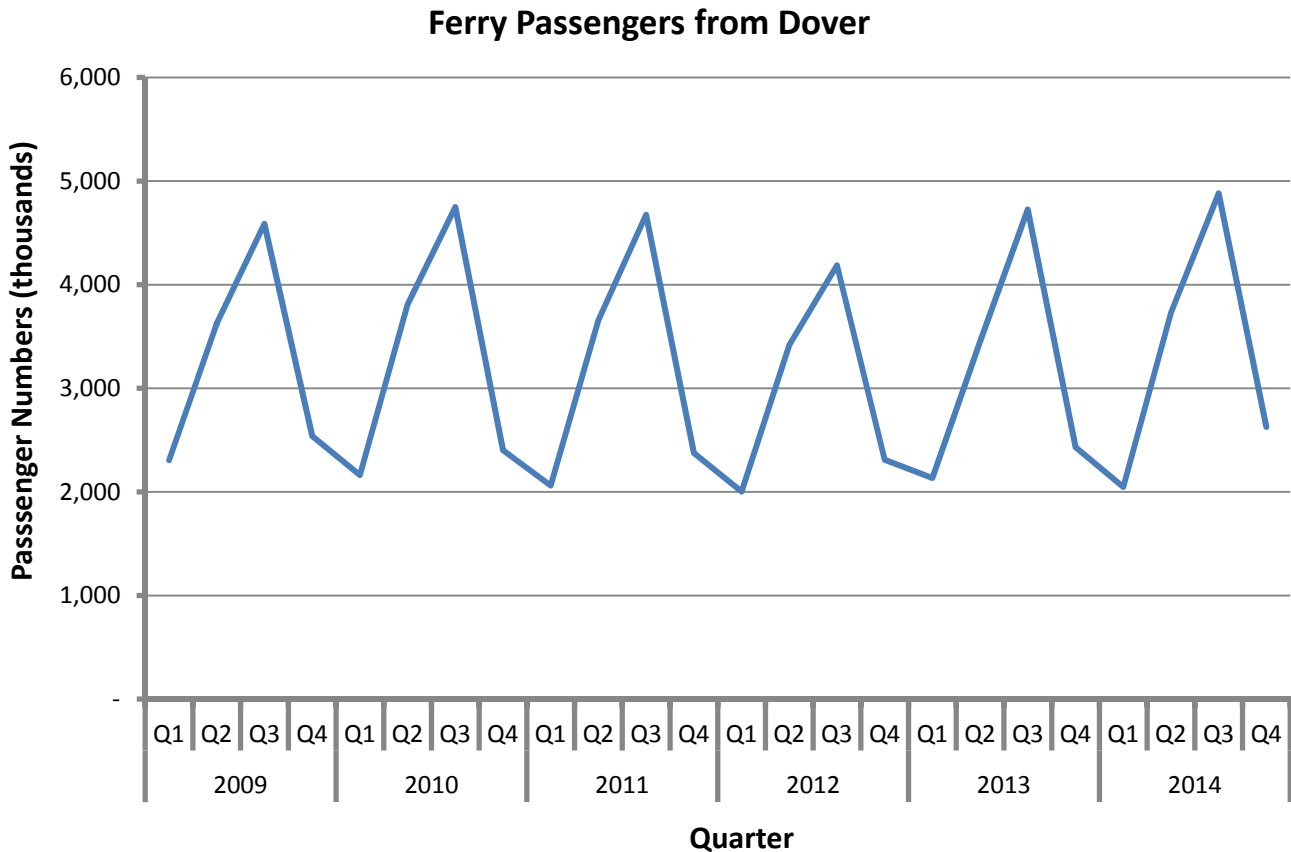
The data was from <https://www.gov.uk/government/statistical-data-sets/spas01-uk-international-sea-passengers> set SPAS0105.

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

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.

Trend

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

Lead in with what you see, and only then move to numbers.

Running a 5 point running average, to remove the seasonal effects, and then finding the trend line for that gave me the graph on the next page.

The gradient for the overall trend is -5.6 , so there has been on average a drop of 5,600 passengers each quarter, or about 22,000 less each year. In the context of 12,000,000 passengers each year it is a very small percentage change (less than 0.2% in fact).

Always put numbers into the context.

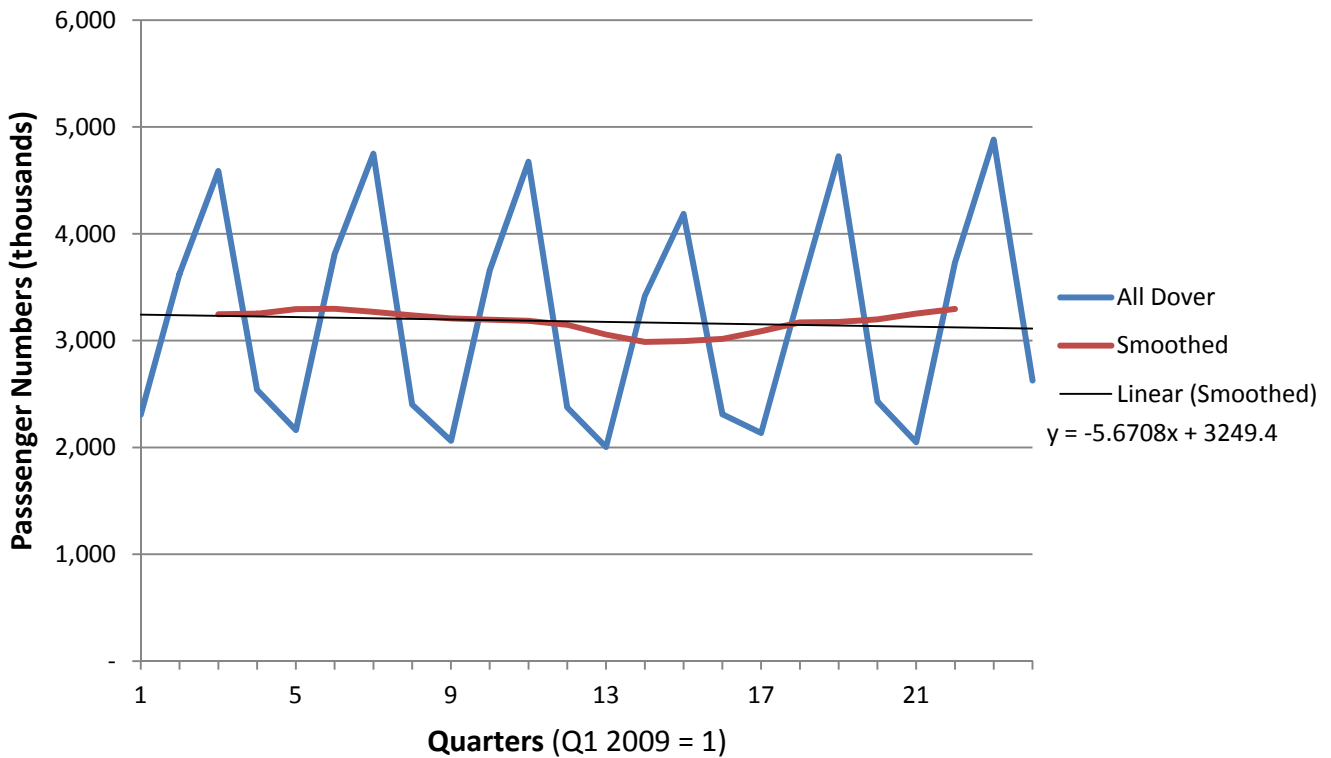
There is one low year of passengers in 2012. As this is in the second half of the time period this has had the effect of dragging down the overall trend, so the trend might be just the result of one very bad year.

Excellence level thinking : if you don't check what the graph shows against the numbers it is easy to miss comments like this. The trend doesn't look down, so why is it?

My research shows that the 10 year trend in ferry passengers from England is slowly downwards in general. This is consistent with the trend from my graph.

There is a result. Is it what we would expect from our research of the topic?

Trend for Ferry Passengers from Dover



Seasonal Pattern

There is a very strong and consistent seasonal pattern to the data, with each year showing a low in Quarter 1, a much higher result in Quarter 2, an even higher result in Quarter 3 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.

I have averaged the difference between the actual values and the trend line for each quarter. The results are below (in thousands):

Q1	-1,068
Q2	434
Q3	1,459
Q4	-723

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.

Explain the meaning of your calculations.

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.

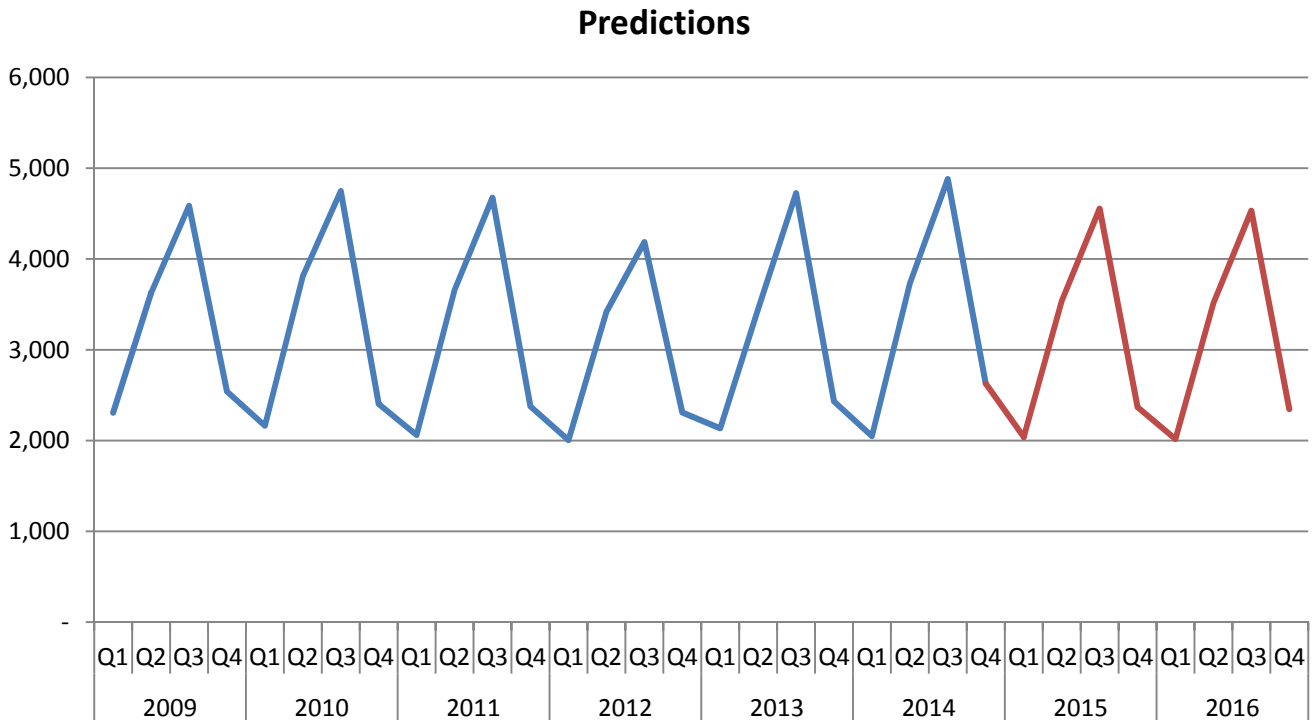
School holidays in England are July and August mostly, that is Quarter 3, which could explain why Quarter 3 is consistently the peak. There is also a school break in May and people who want to travel in good weather avoiding the rush will largely travel in Quarter 2.

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

Predictions

I took the deseasonalised trend calculated and extrapolated it out for eight more quarters, then added the seasonal averages back in to give predictions. The results are shown below.

You must make predictions – it is a compulsory part of the standard. If you cannot make exact numerical ones, then give them in a range of likely values based on visual inspection and the visible pattern.



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.

Unless there are major changes in the circumstances around the ferries from Dover it seems likely that the low quarter predictions (Q1 and Q4) are likely to be accurate. There is very little variation in these quarters for the previous years.

The higher quarters (Q2 and Q3) are likely to be less reliable, because there is much more variation in these quarters in the past. If 2012 was indeed an outlier in its low value then the predictions are likely to be too low.

Discuss the errors.

My research shows little reason to think there are major changes in the near future for the ferries. The Channel Tunnel, which is likely a main reason for the slow downward trend, is 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, so moving them towards ferries, 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.

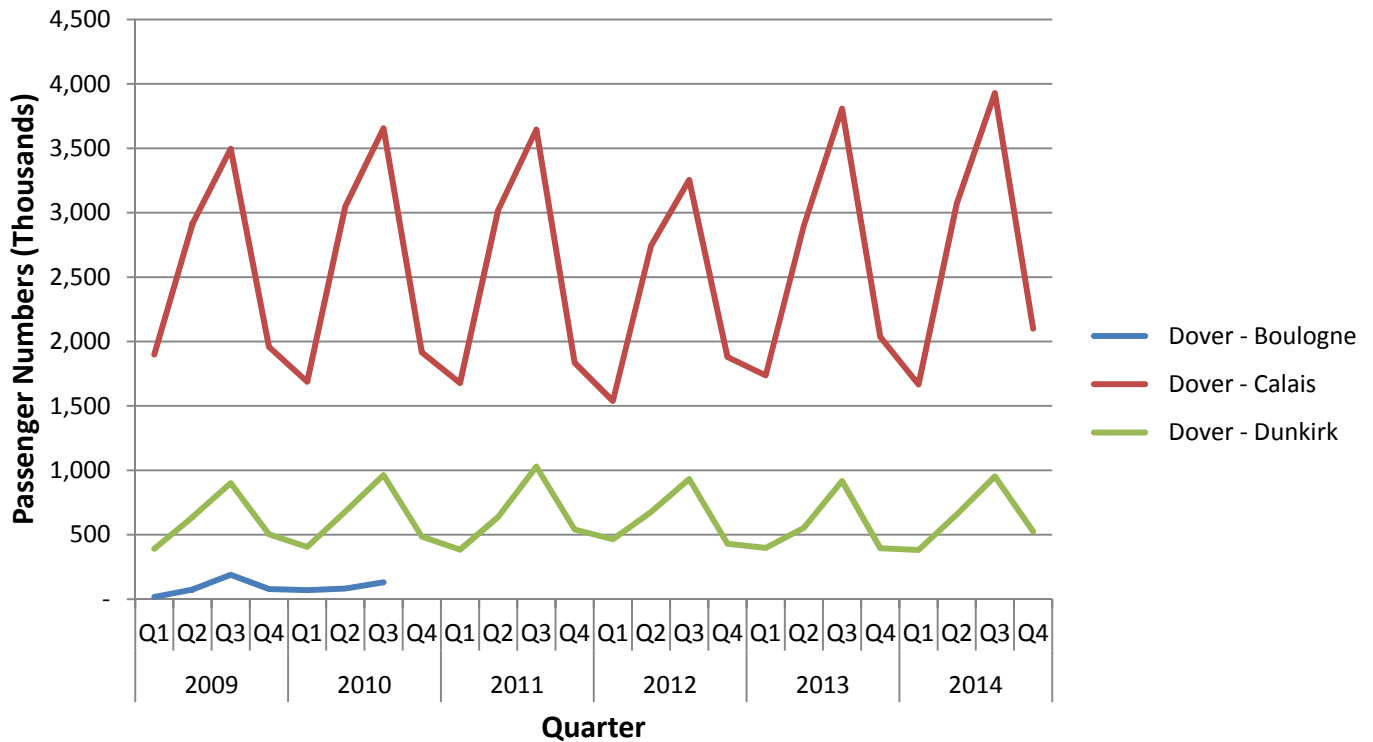
Discuss the likely factors that might affect your predictions, but keep it brief. A few examples of the major issues is sufficient.

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 port:

Dover Ferries by Destination



Travel to Boulogne stopped in late 2010 but was never very important anyway.

It can be seen that the unusually low 2012 values are entirely due to lower travel to Calais, as Dunkirk is unaffected. 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!

The long term trends are also quite different: Calais is actually increasing (slowly) and Dunkirk is dropping by about the same amount. Much of the drop in the overall numbers might be largely the result of the Dunkirk route being dropped.

This suggests that in the future the trend is likely to be quite flat. Perhaps even upwards, as the last three years have seen the Calais route increase faster than the Dunkirk route is falling. However, since businesses are probably more worried by predictions that are too high rather than too low, I think it best to stick with the very small negative overall trend.

While you might do a quick numerical analysis, there is little point showing all the figures like you do with the main analysis.

Other ideas that could have been explored 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 this is about finding some extra interesting feature or some explanation for the main analysis, this is not about repeating the whole process.

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 which is a tiny downward 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.

I would predict then that in the first quarters of 2015 and 2016 that about 2,000,000 people would take the ferries, in the second quarter between 3 and 4 million, in the third quarter between 3 and 4 million, and in the fourth quarter between 2 and 2.5 million.

Don't oversell your predictions – tend to err on the side of caution.

The overall report here is 7 pages, but that is made longer by the annotations. This is more than sufficient: precise explanations with clear graphs is more important than length.