

# H+A Ch. 7: Exponential Smoothing

[Code ▾](#)

*Heather Geiger*

## Libraries

Load all libraries needed to complete this analysis.

[Code](#)

## H+A Question 7.5

### Question

Date set **books** contains the daily sales of paperback and hardcover books at the same store. The task is to forecast the next four days' sales for paperback and hardcover books.

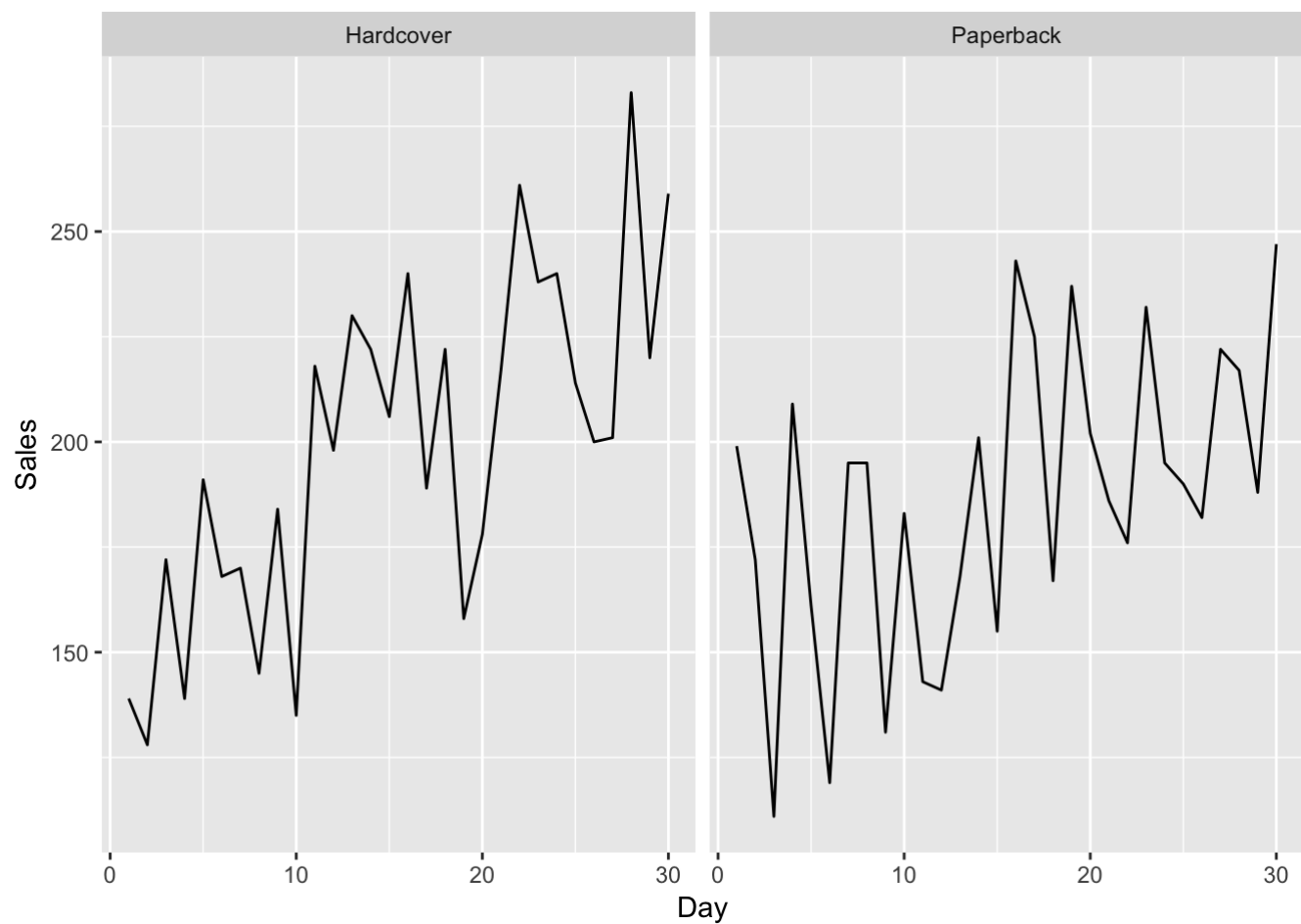
1. Plot the series and discuss the main features of the data.
2. Use the **ses()** function to forecast each series, and plot the forecasts.
3. Compute the RMSE values for the training data in each case.

### Answer

#### Part 1 - Exploratory analysis

Start with a simple line plot.

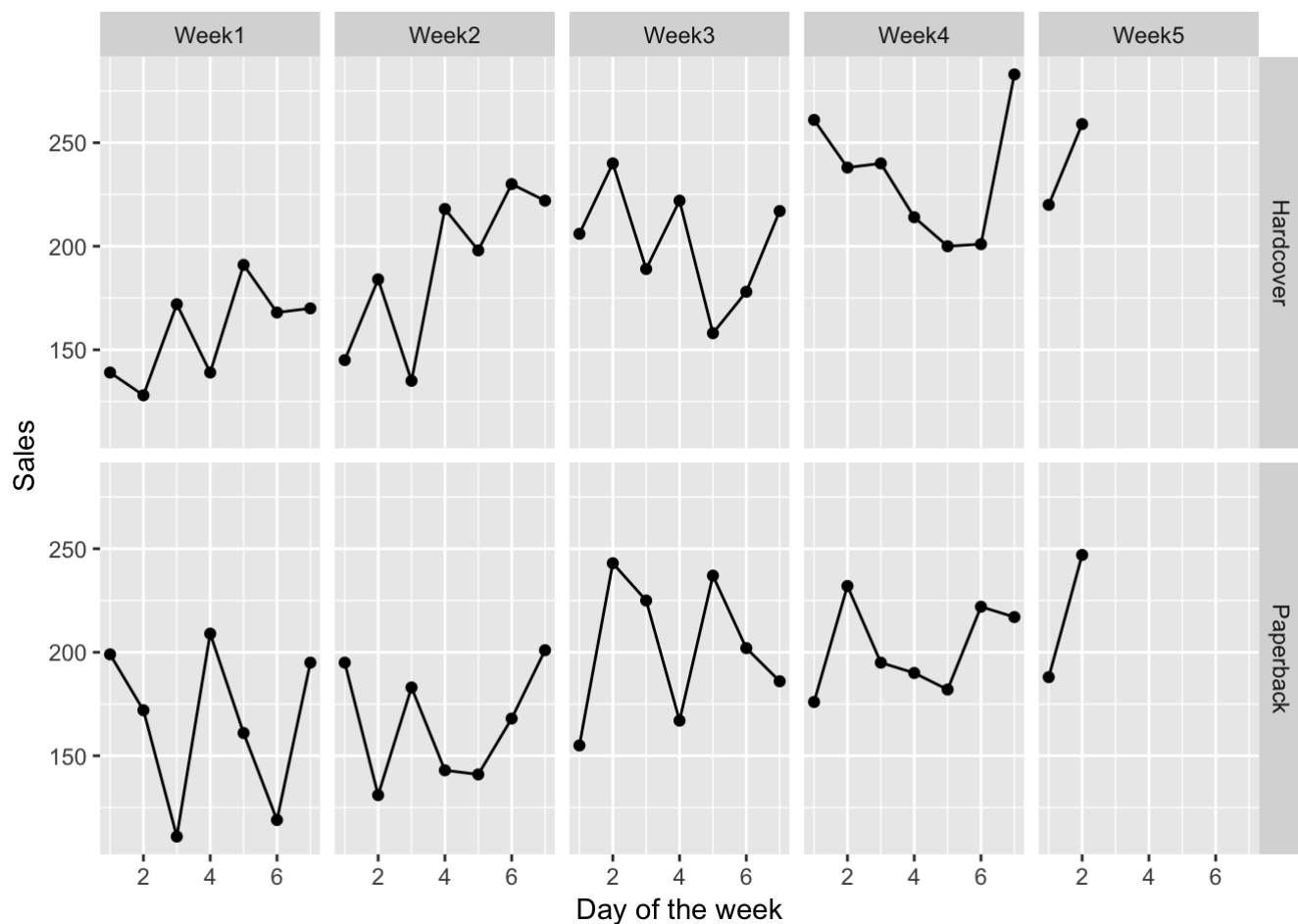
[Code](#)



It feels like there could potentially be a day of the week pattern here, but it's a bit hard to tell from this plot.

Let's plot each set of 7 days in its own panel.

[Code](#)



We don't know here which actual day of the week from Sunday to Saturday each day of the week is, but we could infer if there were a weekday vs. weekend pattern if we saw two consecutive days consistently high/low.

Or if there were one particular day of the week with higher sales for whatever reason, we would be able to see it from this plot.

Here, it really does not seem like there is any pattern within what day of the week it is.

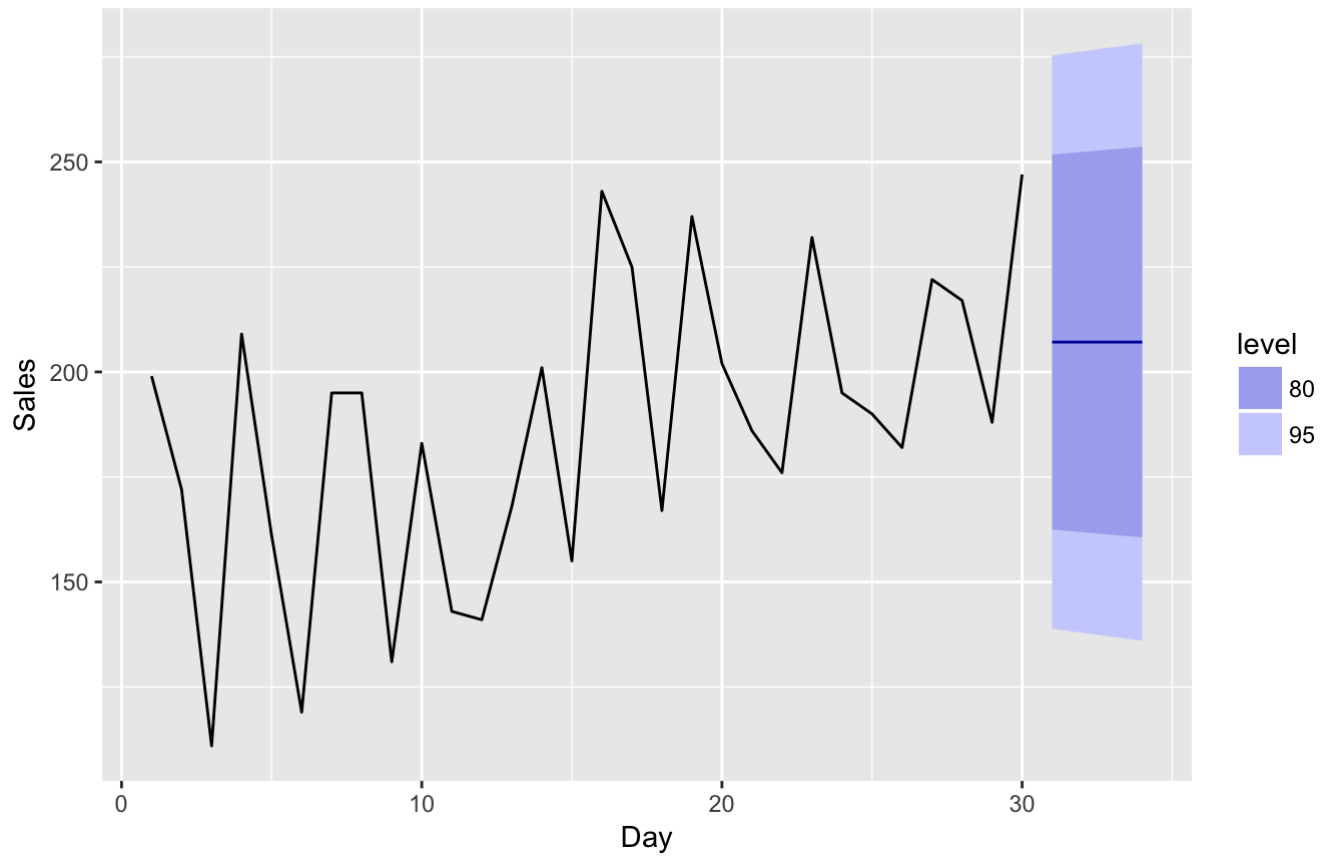
As for trend, it seems like there could possibly be an increasing trend, though it's admittedly a bit hard to tell with such strong daily fluctuations. For our forecasting, we'll move forward based on the assumption that there is no trend and just run simple exponential smoothing.

## Part 2 - Forecasting with ses

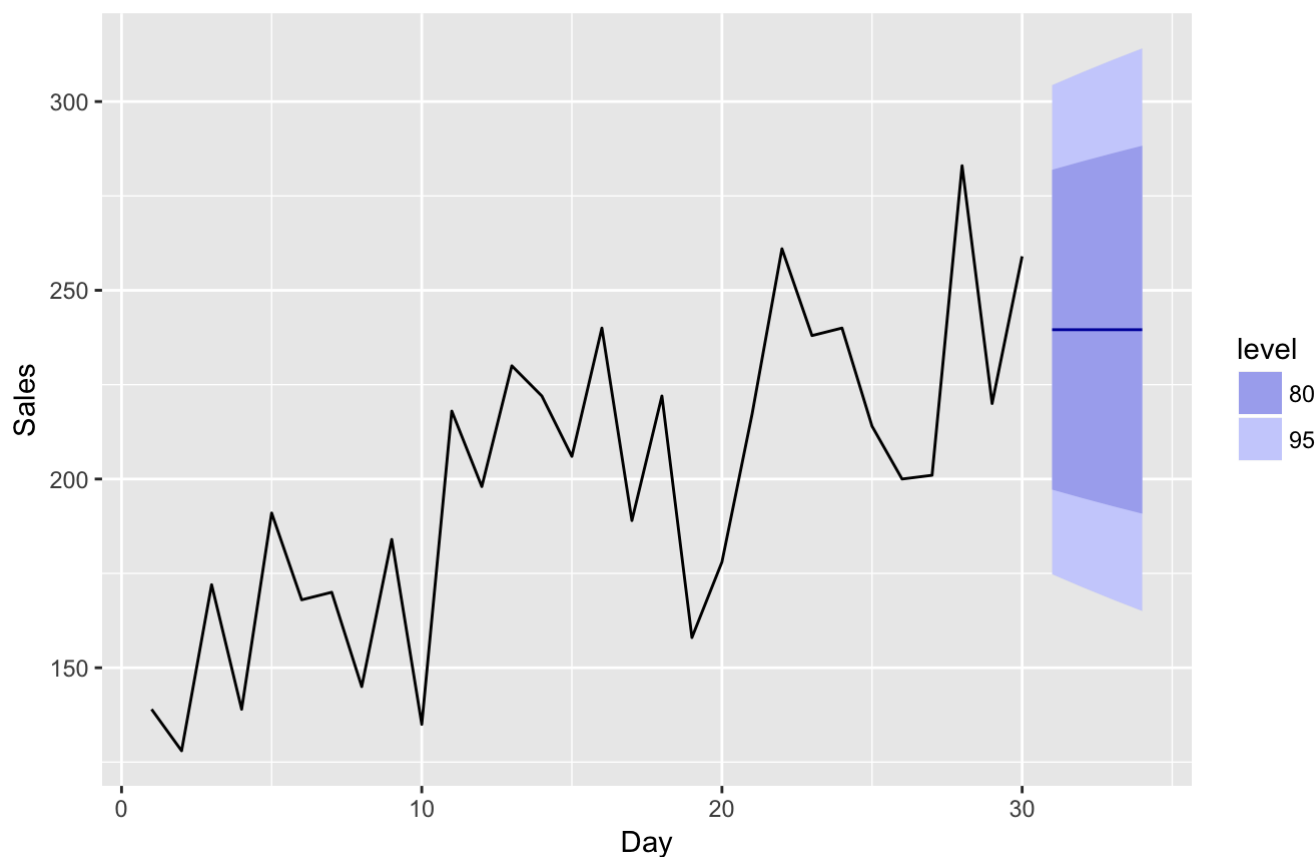
Use the ses function to forecast.

Code

# Paperback sales from past 30 days + forecast through next 4 days (with confidence intervals)

[Code](#)

## Hardcover sales from past 30 days + forecast through next 4 days (with confidence intervals)



Print the numbers for each.

```
## [1] "Paperback forecast:"
```

Code

```
##      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 31      207.1097 162.4882 251.7311 138.8670 275.3523
## 32      207.1097 161.8589 252.3604 137.9046 276.3147
## 33      207.1097 161.2382 252.9811 136.9554 277.2639
## 34      207.1097 160.6259 253.5935 136.0188 278.2005
```

```
## [1] "Hardcover forecast:"
```

Code

```
##      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 31      239.5601 197.2026 281.9176 174.7799 304.3403
## 32      239.5601 194.9788 284.1414 171.3788 307.7414
## 33      239.5601 192.8607 286.2595 168.1396 310.9806
## 34      239.5601 190.8347 288.2855 165.0410 314.0792
```

## Part 3 - Compute RMSE values for the training data in each case.

The MSE (mean squared error) is already included in “mse” within the “model” from the forecast object.

Just take the square root of this to get the RSME (root mean squared error).

```
## [1] "RMSE paperback forecast:"
```

[Code](#)

```
## [1] 33.63769
```

```
## [1] "RMSE hardcover forecast:"
```

[Code](#)

```
## [1] 31.93101
```

The RSME can be thought of as something similar to the standard deviation for the residuals. Basically, it's a way to measure the spread of how much the actual values deviate from the predictions of the model within the training data.

Here, both paperback and hardcover forecasts have relatively similar RSME values of ~32-34. If the residuals are normally distributed, this means we would expect most predictions to fall within +/- 32-34 books vs. the actual value. While the vast majority would fall within +/- 64-68 books.

## H + A Question 7.6

### Question

This question uses the same data as question 7.5.

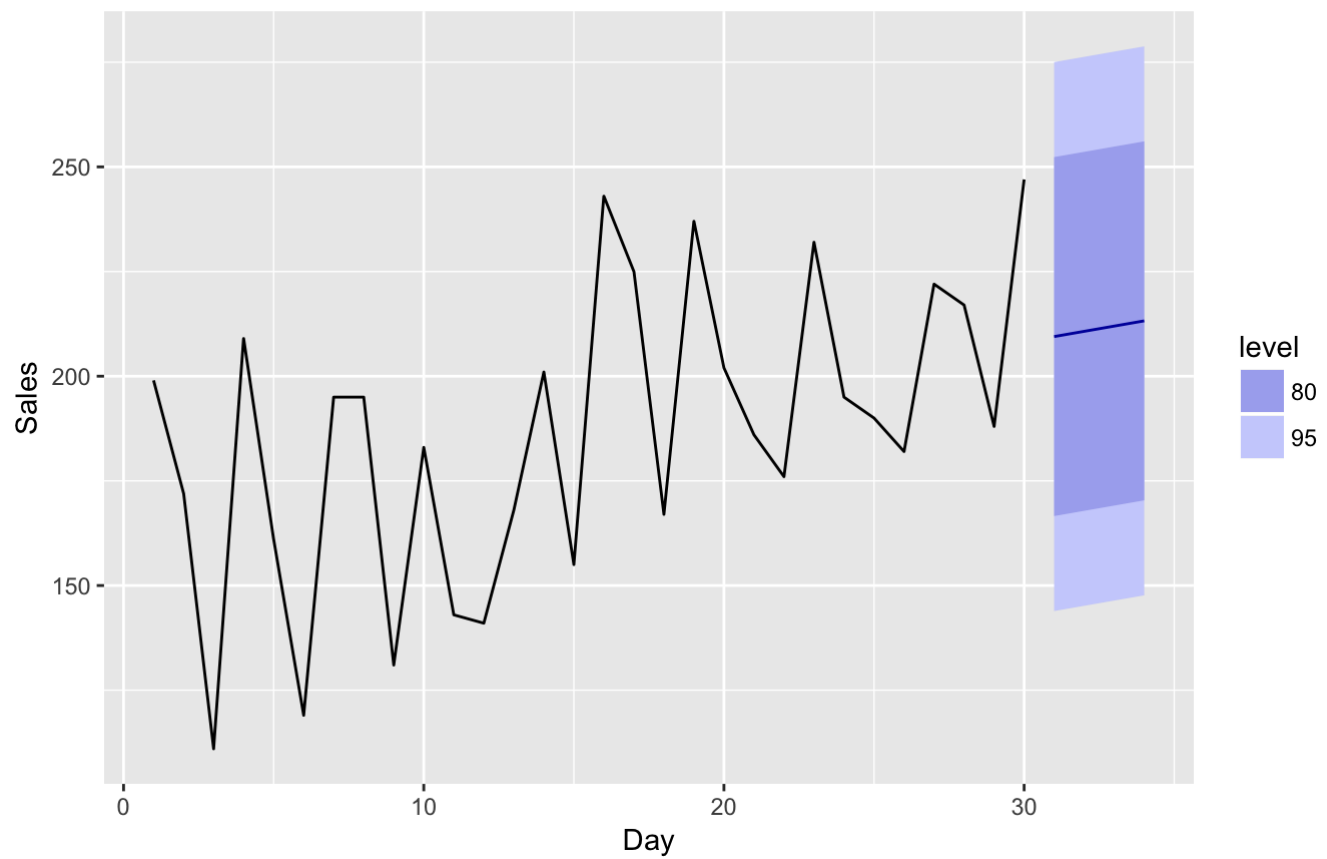
1. Now apply Holt's linear method to the **paperback** and **hardcover** series and compute four-day forecasts in each case.
2. Compare the RMSE measures of Holt's method for the two series to those of simple exponential smoothing in the previous question. (Remember that Holt's method is using one more parameter than SES.) Discuss the merits of the two forecasting methods for these data sets.
3. Compare the forecasts for the two series using both methods. Which do you think is best?
4. Calculate a 95% prediction interval for the first forecast for each series, using the RMSE values and assuming normal errors. Compare your intervals with those produced using **ses** and **holt**.

### Answers

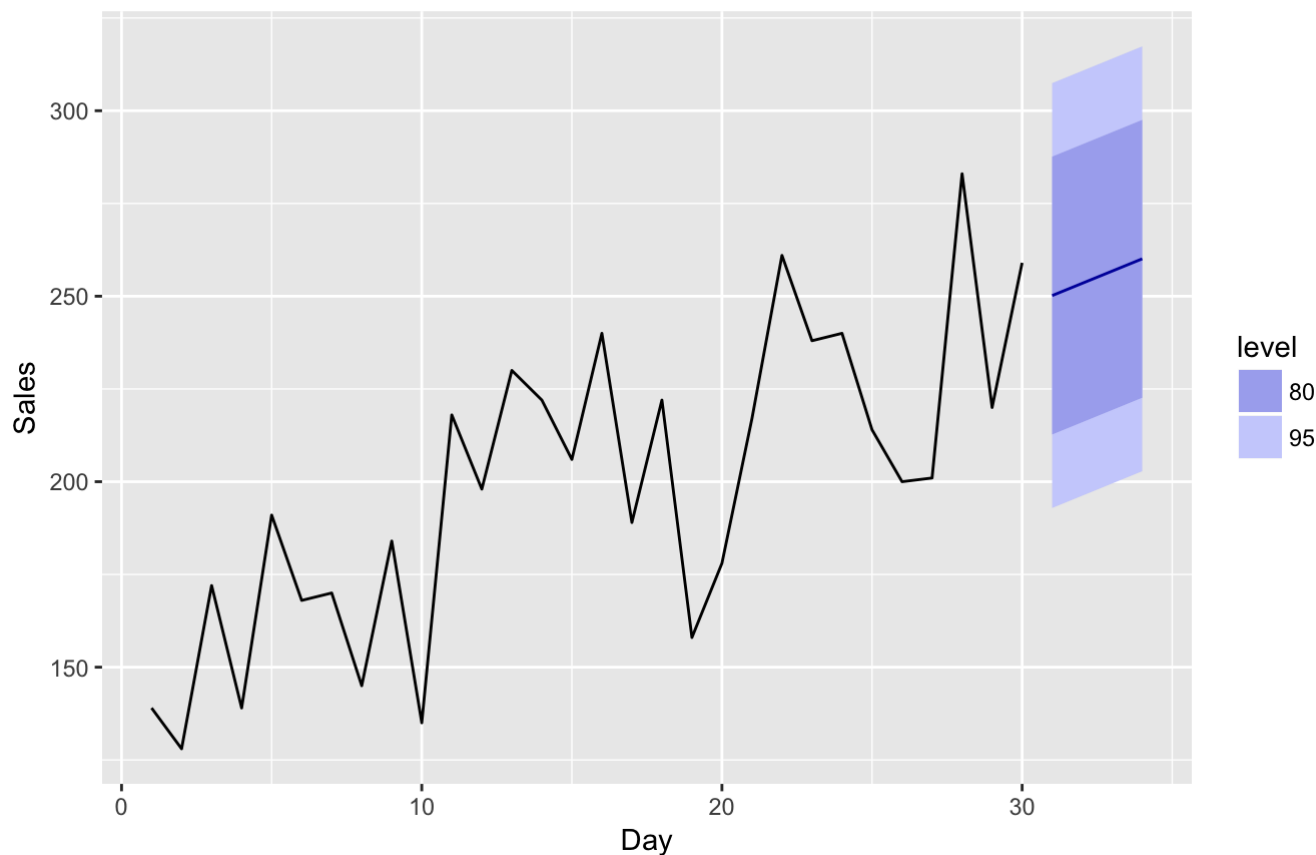
**Part 1 - Apply Holt's linear method to both series and compute four-day forecasts in each case.**

[Code](#)

# Paperback sales from past 30 days + Holt's linear trend forecast through next 4 days

[Code](#)

## Hardcover sales from past 30 days + Holt's linear trend forecast through next 4 days



```
## [1] "Paperback forecast using Holt's linear trend:"
```

[Code](#)

```
##      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 31      209.4668 166.6035 252.3301 143.9130 275.0205
## 32      210.7177 167.8544 253.5811 145.1640 276.2715
## 33      211.9687 169.1054 254.8320 146.4149 277.5225
## 34      213.2197 170.3564 256.0830 147.6659 278.7735
```

```
## [1] "Hardcover forecast using Holt's linear trend:"
```

[Code](#)

```
##      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 31      250.1739 212.7390 287.6087 192.9222 307.4256
## 32      253.4765 216.0416 290.9113 196.2248 310.7282
## 33      256.7791 219.3442 294.2140 199.5274 314.0308
## 34      260.0817 222.6468 297.5166 202.8300 317.3334
```

## Part 2 - Compare RSME measures vs. simple exponential smoothing.

First, calculate RSME for Holt's forecasts.



```
## [1] "RSME for Holt's on paperback"
```

[Code](#)

```
## [1] 31.13692
```

```
## [1] "RMSE for Holt's on hardcover"
```

[Code](#)

```
## [1] 27.19358
```

The RMSE error values using Holt's linear trend is lower for both series than using simple exponential smoothing.

Assuming both methods have a mean of the residuals around 0, this would mean that we should prefer Holt's linear trend, since its predictions show less deviation from being perfectly accurate.

However, this is a big assumption! Let's check if it's true.

```
## [1] "Means of residuals for SES, then Holt's, for paperback series:"
```

[Code](#)

```
## [1] 7.175981
```

[Code](#)

```
## [1] -3.717178
```

```
## [1] "Means of residuals for SES, then Holt's, for hardcover series:"
```

[Code](#)

```
## [1] 9.166735
```

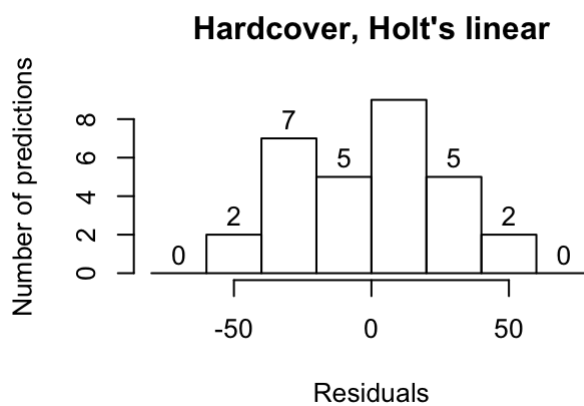
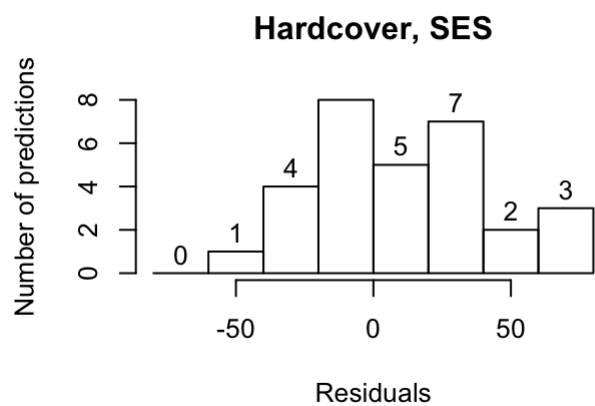
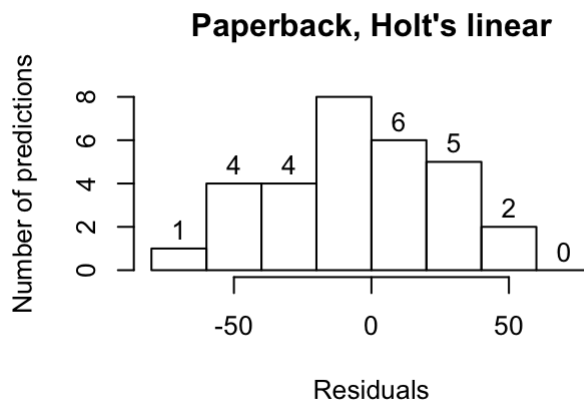
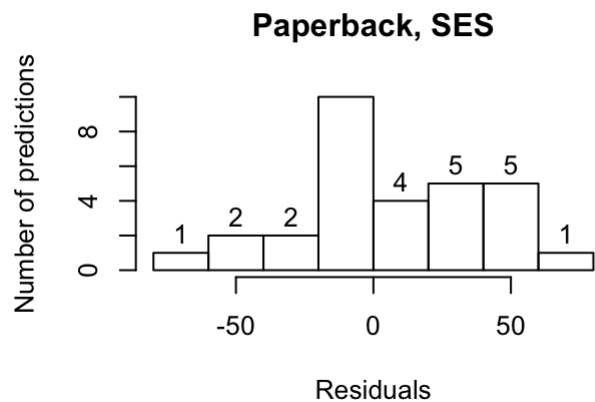
[Code](#)

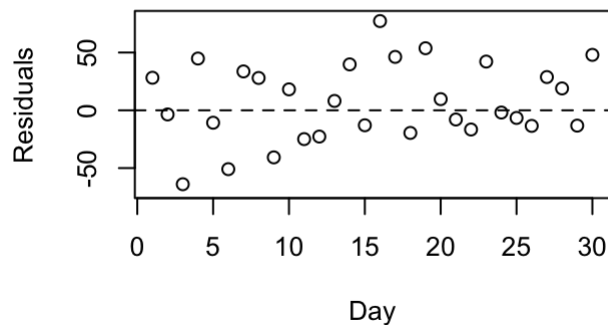
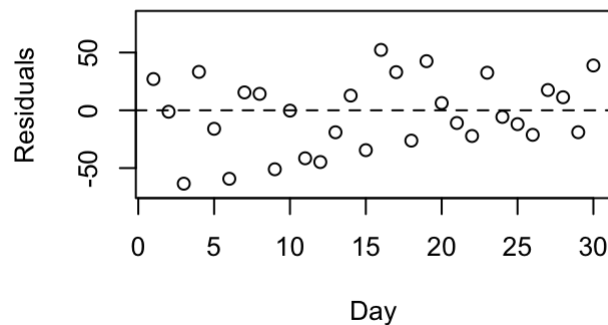
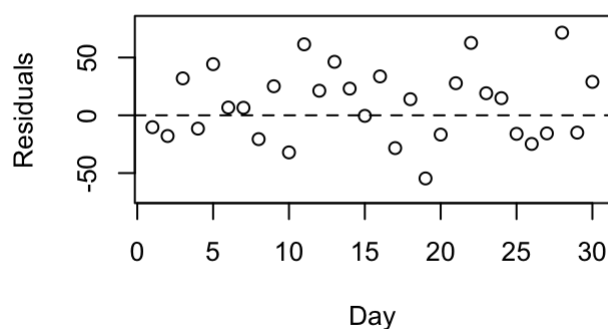
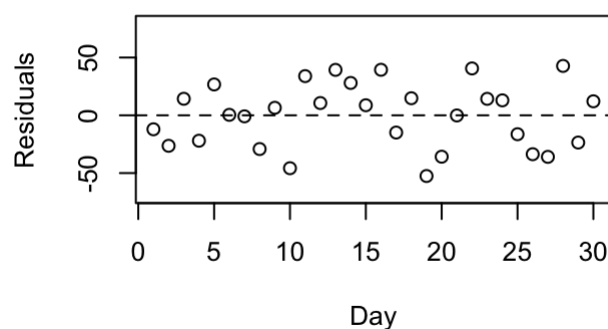
```
## [1] -0.1357882
```

Looks like the mean of the residuals using Holt's is actually closer to 0 than for SES, so that's another check in favor of Holt's.

Finally, let's visualize the distribution of residuals as well as plot time vs. residuals.

[Code](#)

[Code](#)

**Paperback, SES****Paperback, Holt's linear****Hardcover, SES****Hardcover, Holt's linear**

The distributions of residuals look pretty similar. It's hard to tell with only 30 data points, but they all seem relatively close to normal.

There does not seem to be any major bias across time for predictions to be in any one direction.

Therefore with these other factors being equal and RSME being lower using Holt's linear trend, it seems that Holt's linear trend has the clear advantage if we consider only performance within the training data.

## Part 3 - Compare the forecasts for the two series using both methods. Which do you think is best?

We decided in part 2 that Holt's linear trend would definitely be considered best if we looked only at performance within the training data.

Now, we are asked to also compare them using the forecasts into the future.

We saw above that SES gives the same forecast value for all four days into the future, with only the prediction intervals changing over time. Whereas Holt's linear trend gives a different (in this case increasing) value for each point further out in time.

Paperback SES gave a prediction of 207.1097 books for all four days.

Meanwhile Holt's method gave a prediction pretty close to the prediction from SES (209.4668 books) on the first day. But by the last day its prediction was decently higher (213.2197 books).

Similar idea for the hardcover series (239.5601 from SES, 250.1739 to 260.0817 from Holt's linear method).

Let's look at how many days had sales higher than these different predictions within each series.

```
## [1] "Days with sales higher than SES prediction, then first day Holt's prediction, then last day Holt's prediction, paperback series:"
```

Code

```
## [1] 8
```

Code

```
## [1] 7
```

Code

```
## [1] 7
```

```
## [1] "Same for hardcover series:"
```

Code

```
## [1] 5
```

Code

```
## [1] 3
```

Code

```
## [1] 2
```

We find that both SES and Holt's method predict future sales that are higher than most sales of the past 30 days, for both series. So while the predictions using Holt's method are even higher than for SES, it's not like they are that much more out of the realm than the predictions from SES.

With this information combined with the better performance within the training data, I would continue to say that Holt's linear trend is the better choice over SES for forecasting from this data.

## Part 4 - Calculate prediction intervals and compare to intervals from SES and Holt functions.

First, let's review the intervals from SES and Holt functions.

Prediction interval for first forecast from SES, paperback series: 138.8670-275.3523 (range ~136)

Holt's linear trend, paperback series: 143.9130-275.0205 (range ~131)

SES, hardcover series: 174.7799-304.3403 (range ~130)

Holt's linear trend, hardcover series: 192.9222-307.4256 (range ~115)

Now, compare to what we would get from calculating it ourselves using RMSE and assumption of a normal distribution.

Code

```
## [1] "Intervals for SES then Holt's, paperback series:"
```

[Code](#)

```
## [1] 141.1810 273.0383
```

[Code](#)

```
## [1] 148.4395 270.4940
```

```
## [1] "Intervals for SES then Holt's, hardcover series:"
```

[Code](#)

```
## [1] 176.9765 302.1437
```

[Code](#)

```
## [1] 196.8754 303.4723
```

To summarize again:

SES, paperback: 138.8670-275.3523 from function vs. 141.1810-273.0383 calculated

Holt's, paperback: 143.9130-275.0205 vs. 148.4395-270.4940

SES, hardcover: 174.7799-304.3403 vs. 176.9765-302.1437

Holt's, hardcover: 192.9222-307.4256 vs. 196.8754-303.4723

All of our calculated intervals are very slightly narrower than those given by the forecasting functions.

Overall though, they are quite similar.

## H+A Question 7.10

### Question

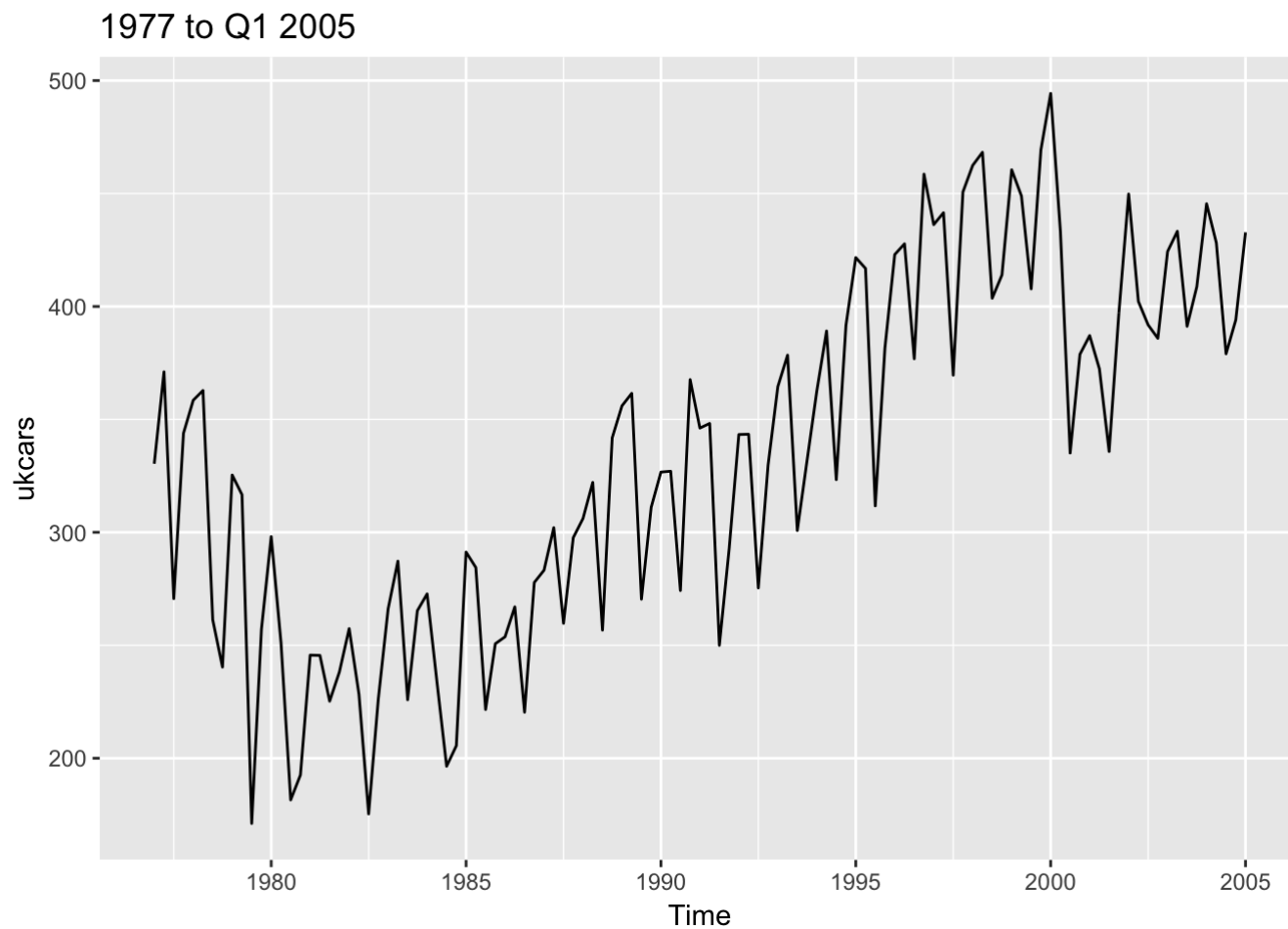
For this exercise use data set **ukcars**, the quarterly UK passenger vehicle production data from 1977Q1–2005Q1.

1. Plot the data and describe the main features of the series.
2. Decompose the series using STL and obtain the seasonally adjusted data.
3. Forecast the next two years of the series using an additive damped trend method applied to the seasonally adjusted data. (This can be done in one step using `stlf()` with arguments `etsmodel="AAN"`, `damped=TRUE`).
4. Forecast the next two years of the series using Holt's linear method applied to the seasonally adjusted data (as before but with `damped=FALSE`).
5. Now use `ets()` to choose a seasonal model for the data.
6. Compare the RMSE of the ETS model with the RMSE of the models you obtained using STL decompositions. Which gives the better in-sample fits?
7. Compare the forecasts from the three approaches? Which seems most reasonable?
8. Check the residuals of your preferred model.

# Answers

## Part 1 - Exploratory analysis

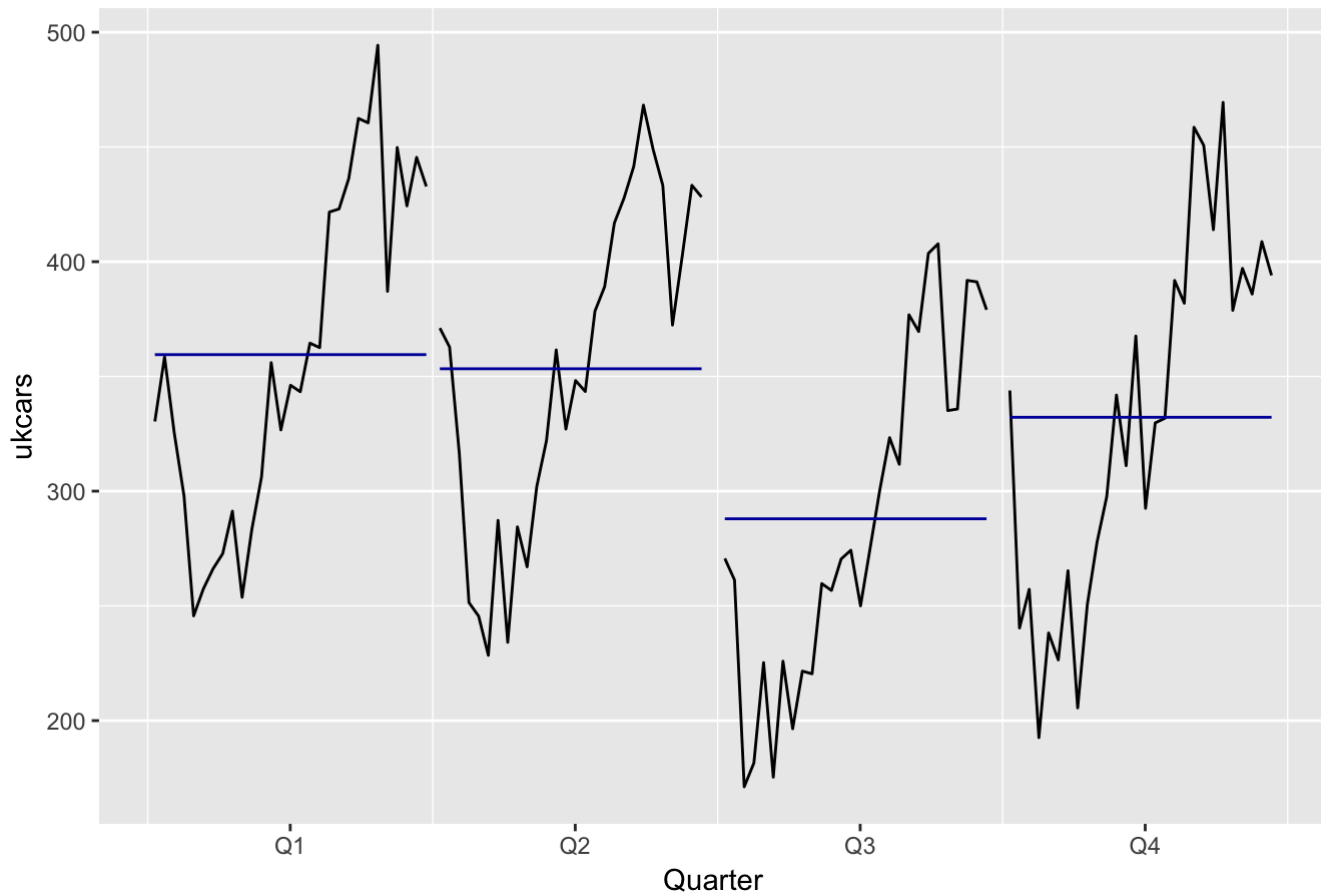
Start with a simple line plot.

[Code](#)

Also look at each quarter separately.

[Code](#)

## Yearly performance from 1977 to 2004/2005, separated by quarter



There is a clear seasonal component to the data, with lower sales on average in Q3.

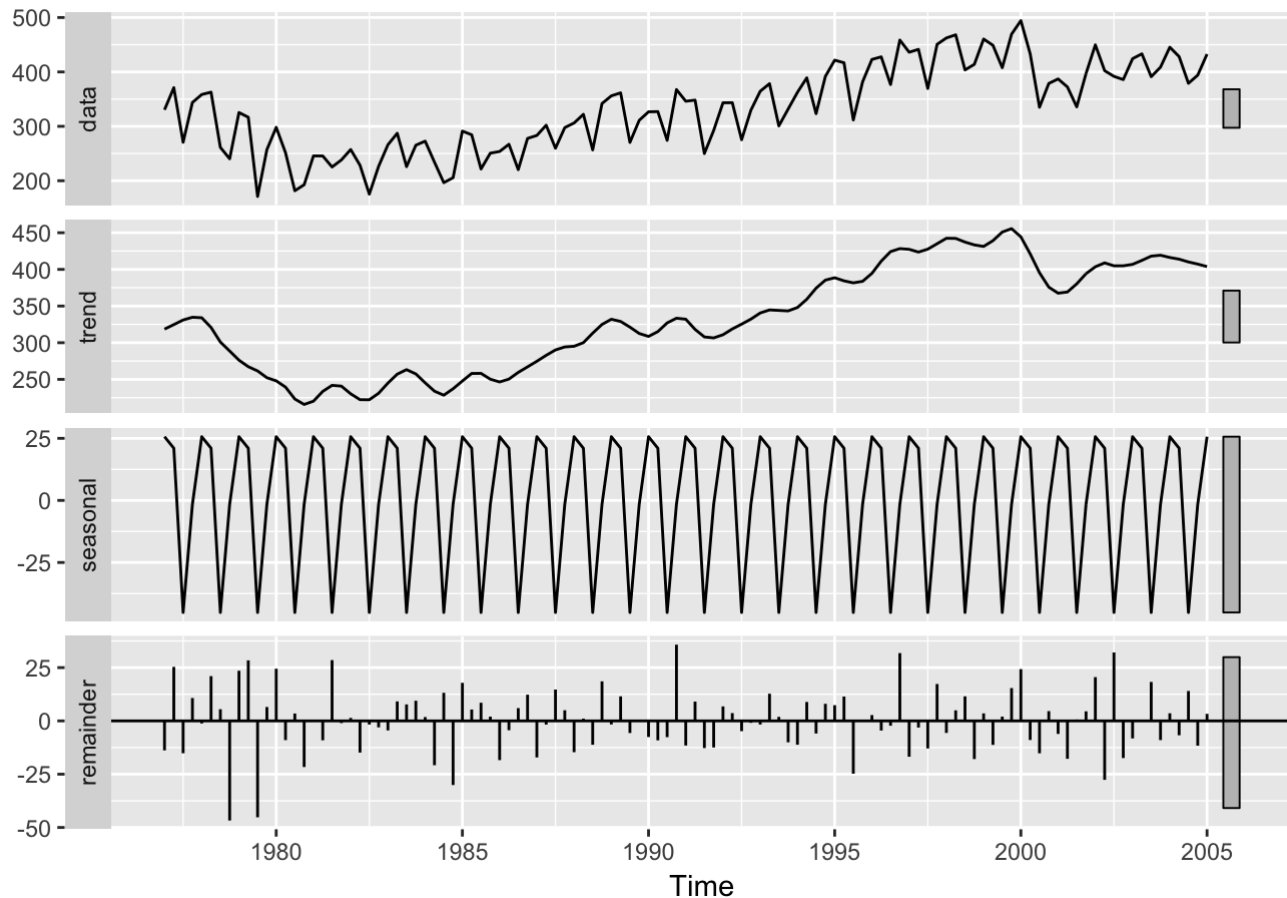
There also appears to be an increasing trend, at least from around 1981-2000 or so.

## Part 2 - STL decomposition to seasonally adjust data

Let's use the **stl** function to do this, and plot the result.

[Code](#)

### STL decomposition of ukcars



These results look pretty sensible. We can clearly see the dips in Q3 in the seasonal component, and the increasing trend throughout most of the time span in the trend component.

## Part 3 - Forecast using additive damped trend.

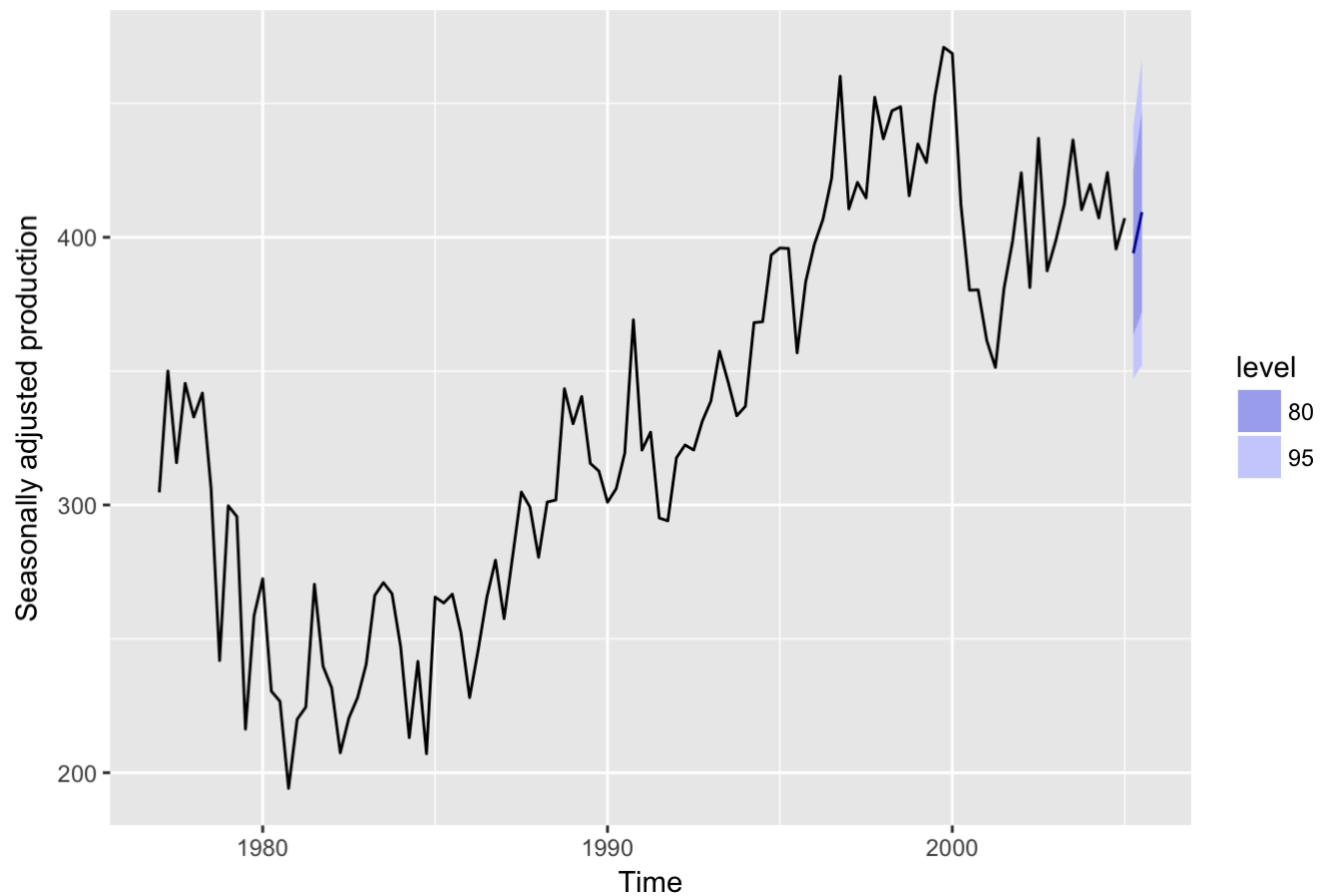
As described in the question, use the function `stlf()` to do this.

Then, plot the seasonally adjusted data and corresponding forecast.

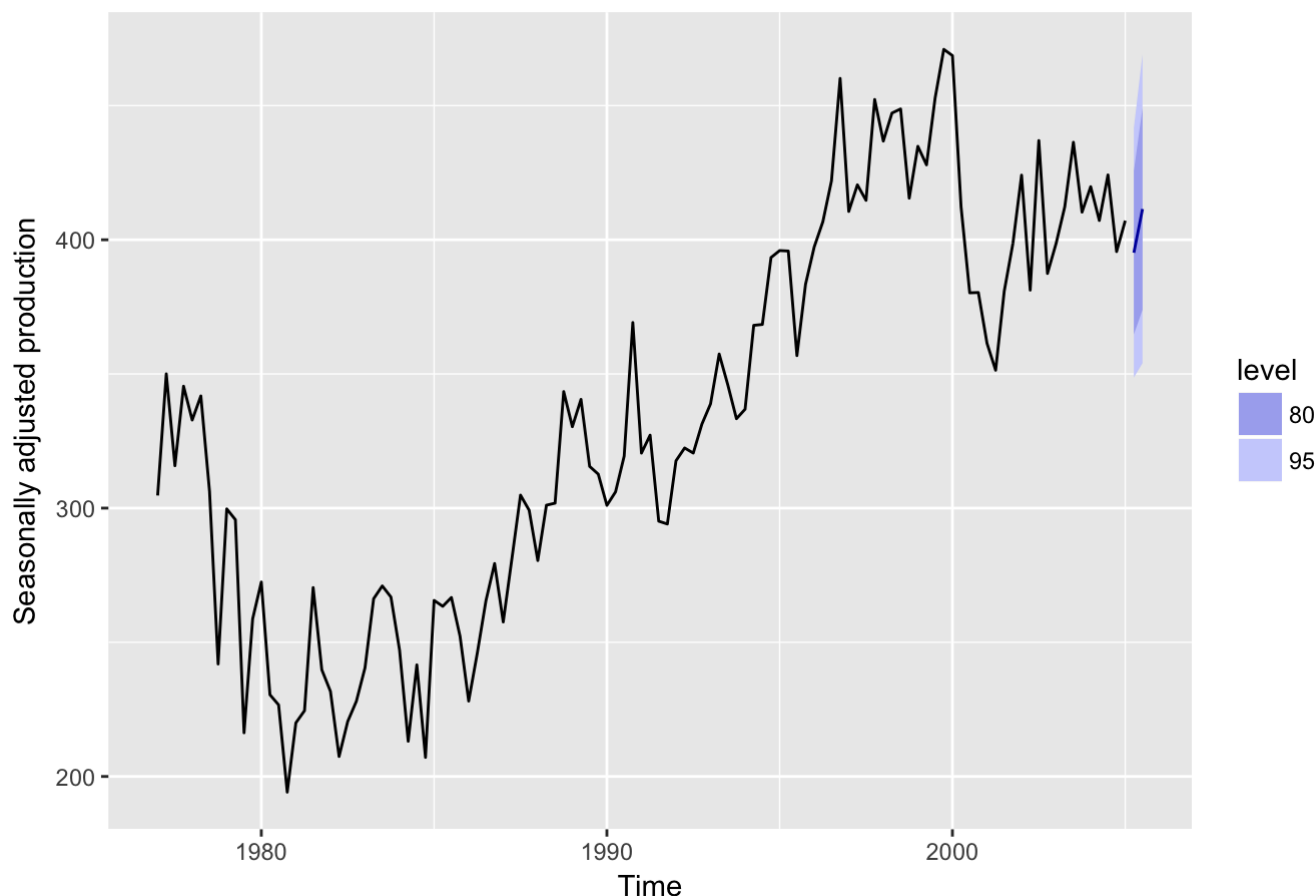
[Code](#)



## Forecasts from STL + ETS(A,Ad,N)

**Part 4 - Forecast using additive trend, NOT damped.**[Code](#)

## Forecasts from STL + ETS(A,A,N)



## Part 5 - Choose seasonal model for the data using the ets function.

Code

```
## [1] "A"      "N"      "A"      "FALSE"
```

The third letter denotes the season type. Here we get “A”, aka an additive seasonal model.

## Part 6 - Compare RSME from these models.

```
## [1] "RSME damped trend after STL decomposition, non-damped after STL decomposition, and ETS automated model selection:"
```

Code

```
## [1] 23.32113
```

Code

```
## [1] 23.295
```

Code

```
## [1] 25.23244
```

Based on RSME, both models after STL decomposition perform better on the training data than the one using non-adjusted data and the ETS function.

The non-damped performs slightly better than the damped, but they are really extremely close.

## Part 7 - Compare forecasts from these models.

```
## [1] "Forecasts damped trend after STL decomposition, non-damped after STL decomposition, and ETS automated model selection:"
```

Code

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## 2005 Q2	394.0146	363.4433	424.5858	347.2599	440.7692
## 2005 Q3	409.3791	371.9947	446.7634	352.2047	466.5535

Code

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## 2005 Q2	395.1772	364.7806	425.5738	348.6897	441.6648
## 2005 Q3	411.4591	373.9155	449.0027	354.0411	468.8771

Code

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## 2005 Q2	427.4885	394.2576	460.7195	376.6662	478.3109
## 2005 Q3	361.3329	322.2353	400.4305	301.5383	421.1275

The damped forecast has slightly lower forecast point values than the non-damped, as we would expect.

The ETS model predicts much higher Q2 production and much lower Q3 production than the other two models. However, it is hard to compare directly because the ETS model is based on non-adjusted data, while the others are based on adjusted data.

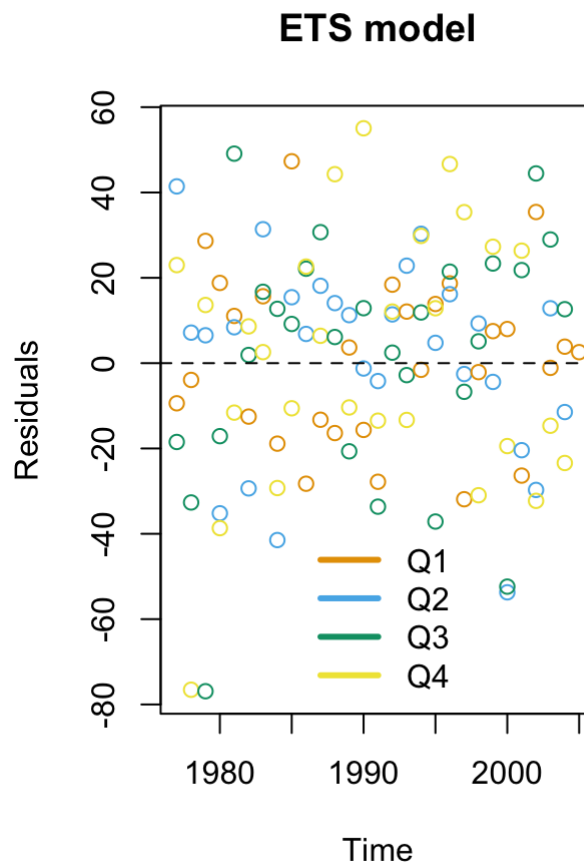
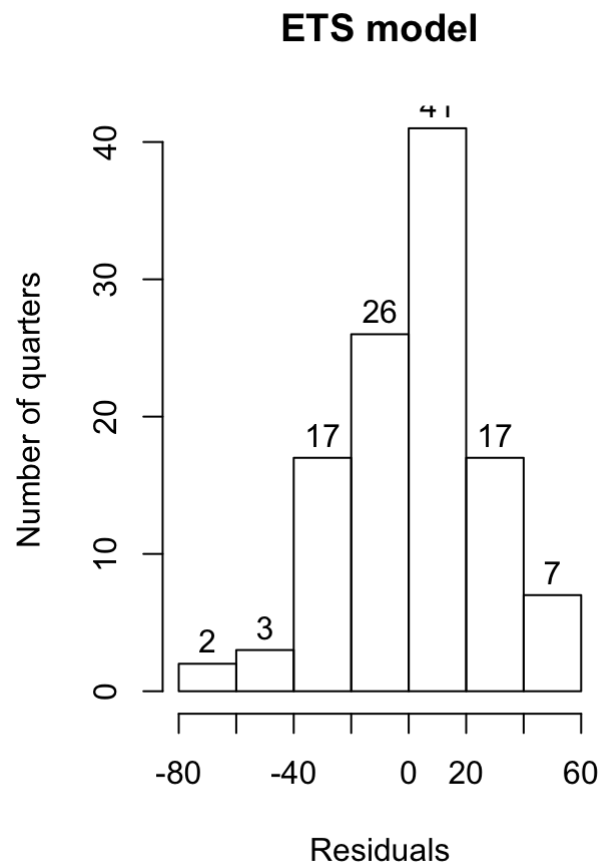
To convert the forecasts after STL decomposition to real-world realistic numbers, we would have to add back in the seasonal component to these forecasts.

Based on this, plus the fact that the RSME was not *that* much worse for the ETS model, I will proceed with the ETS model from here on out.

## Part 8 - Check residuals of preferred model.

Let's make a histogram and plot time vs. residuals for the ETS model.

Code



Residuals appear to be normally distributed and unbiased over time and season.