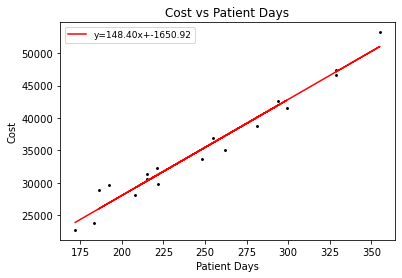
## Side Note: This document just shows my final answers/graphs. The work was all done in Python (Jupyter).

## Question 1: Exercise 2.13

1. Using a linear regression model, develop predictions for costs in periods 1 through 18 based on patient days.

See Python file: q1 data frame.



1. What is the prediction error in period 1? What is the prediction error in period 18?

Period 1: $454.10

Period 18: $1169.58

1. Predict costs when patient days are 200, 300, and 400.

## 200 Days: $800.61

## 300 Days: $1157.88

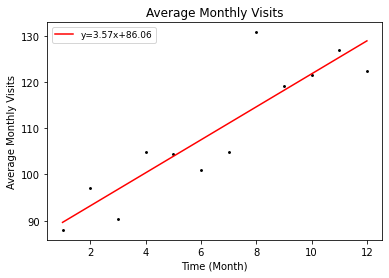
## 400 Days: $1515.15

## Question 2: Exercise 2.19

1. Develop a linear regression prediction model based on the average visits in each month.

(x, y) = (month number, average visits each month)

y = 3.57x + 86.06



1. Predict the visits for January through March of the following year.

January: 90.0 Visits

February: 93.0 Visits

March: 97.0 Visits

1. Develop daily and monthly indexes for urgent care clinic visits.

See Python file.

Daily Index: q2\_transpose data frame

Monthly Index: q2 data frame

1. Use the monthly indexes technique to develop the monthly adjusted predictions for January through March.

See Python file: q2 data frame

## Question 3: Exercise 2.26

1. Predict enrollment in November of year 5 using a five-period moving average.

Predicted Enrollment in November of year 5: 1257.

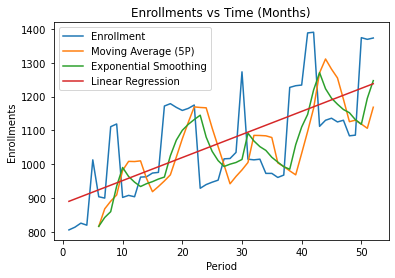
1. Predict enrollment in November of year 5 using the simple exponential smoothing model with α = 0.3.

Predicted Enrollment in November of year 5: 1285.

1. Predict enrollment in November of year 5 using linear regression.

Predicted Enrollment in November of Year 5: 1245.

1. Plot the actual enrollment against the predictions developed in (a) through (c).



1. Calculate MAD and MAPE for the predictions developed in (a) through (c).

Moving Average

* MAD: 128.09
* MAPE: 11.38%

Exponential Smoothing

* MAD: 99.43
* MAPE: 8.81%

Linear Regression

* MAD: 102.27
* MAPE: 9.44%

1. Which prediction method would you recommend? Explain.

I would recommend predicting the enrollments using exponential smoothing because it has the lowest MAD and MAPE, which shows less deviation and percentage errors. The graph shown above indicates that the exponential smoothing line is the closest to the actual enrollment line on the graph. It follows the seasonality the most accurately of the three forecasting methods, whereas the other methods lag a bit behind when following the actual enrollment seasonality.

## Question 4: Exercise 2.27

1. Using linear regression, predict appointments for the first four weeks in January of the following year. Plot the actual appointments against the predicted values.

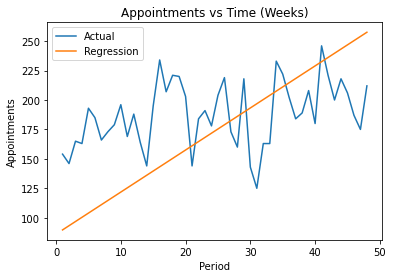
Regression Equation: y = 0.705438558402084\*x + 170.4875886524823

Week 1: 205 Appointments

Week 2: 206 Appointments

Week 3: 206 Appointments

Week 4: 207 Appointments



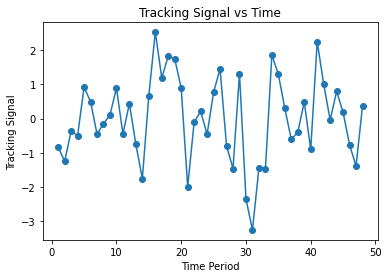
1. Using the prediction results, calculate the MAD and MAPE.

MAD: 20.67

MAPE: 11.50%

1. Compute and graph the tracking signal for the prediction.

See Python file: q4 data frame.



1. What do you conclude about this model?

Regression is not a good for this particular model because there are clear trends, such that a constant regression line does not accurately reflect the number of appointments per time period.

## Question 5: Exercise 2.30

1. Navigate to https://data.cms.gov/. Select the “ACO” category and locate the “Medicare Shared Savings Program Accountable Care Organizations Performance Year 1 Results” data set. Click “Export” and download the file into MS Excel format.

See Python file: q5 data frame.

1. Develop a linear regression model that predicts total expenditures based on the number of assigned beneficiaries.

(x, y) = (total assigned beneficiaries, total expenditures)

y = 11,830.91x – 5,526,996.21

1. Predict expenditures when the number of assigned beneficiaries is 10,000. What is the percentage change in predicted total expenditures when the number of beneficiaries increases from 10,000 to 20,000?

Expenditures when the number of assigned beneficiaries is 10,000: $112782143.19.

Percentage Change in Total Expenditures: 104.90%

1. Predict expenditures when the number of assigned beneficiaries is 50,000. What is the percentage change in predicted total expenditures when the number of beneficiaries increases from 50,000 to 60,000?

Expenditures when the number of assigned beneficiaries is 50,000: $586,018,700.80.

Percentage Change in Total Expenditures: 20.19%

1. Predict expenditures when the number of assigned beneficiaries is 100,000. What is the percentage change in predicted total expenditures when the number of beneficiaries increases from 100,000 to 110,000?

Expenditures when the number of assigned beneficiaries is 100,000: $1177564397.82.

Percentage Change in Total Expenditures: 10.05%

1. What is the predicted number of assigned beneficiaries when total expenditures are $100 million? What is the predicted number of assigned beneficiaries when total expenditures are $500 million?

Expenditure total is $100000000: 8920.0 assigned beneficiaries.

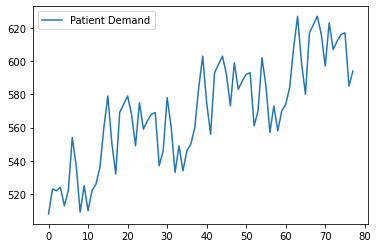
Expenditure total is $500000000: 42729.0 assigned beneficiaries.

1. Based on your analysis, does this model suggest that ACOs are more cost-effective? Support your conclusion.

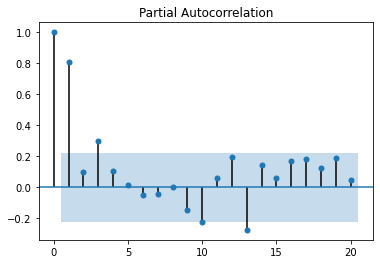
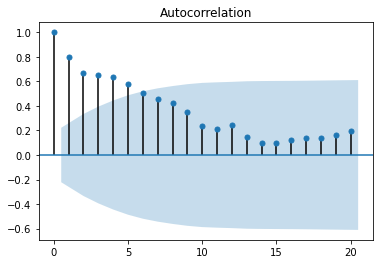
It seems like the percentage changes to the assigned beneficiaries are proportional to the changes in expenditures accordingly. This makes it seem that the ACOs are more cost effective, given that most expenditures increase disproportionately when they take on more assigned beneficiaries.

## Question 6:

To ensure service quality in emergence department, a hospital manager likes to have a strong truth about patient arrival pattern so that he can plan sufficient resources in advance. He is looking for an accurate forecast tool for patient arrival prediction. The historical data is given in the attached Excel file. The last 10 days are used for forecasting model validation.



1. Plot ACF and PACF with lag k = 1 – 20. What do you observe from the historical patient arrivals, e.g, trend, seasonality, noise, etc.?



From these graphs, I have noticed that there is a trend every 30 periods that repeats itself, granted that the appointments increase every 30 periods, but the trend and seasonality are both present in this graph. The spikes look identical in size, except that they have been shifted up in the graph over time.

In the autocorrelation plot, it is clear that the autocorrelation values start to increase again at the 20th point after the initial decrease. This is reflective of the seasonality/trend present in the demand graph. The partial autocorrelation plot also shows an initial decrease followed by a sudden increase in values, which indicates some level of trend and/or seasonality, which is shown in the first plot of the demand over time.

1. Develop your forecast model ARIMA (p, d, q). Elaborate your determination for the best settings.

Visuals are in the Python file.

I first looked at the patient demand graph, and I noticed seasonality and trend, where it repeats every 30 periods. To get the value for (p, d, q), I first noticed that after the third period, the demand direction changed, so I decided to use p=3 so the correlation would be significant for the first 3 values. For d, I used d=1 (difference order) to make the time series stationary, and I used a moving average model of 0 (q=0). This worked out well, as the distribution of the errors were pretty standardized (mean of 0). So I used (3, 1, 0) for values (p, d, q) respectively.

**References**

1. *Analytics and Decision Support in Health Care Operations Management*, 3rd Edition, Yasar A. Ozcan, 2017