Blood Bags Containers Forecast **Dion Chang**

THE SCENARIO: I am working for a blood bank, and we are considering ordering some containers to improve the efficiency of transporting blood bags that are collected from donors. These containers take a 3-days basis to be delivered to the blood bank starting from October 1st, 2018.

There are also various types of containers for each kind of blood samples: Whole Blood, Platelets, and Packed Red Cells.

The blood bank prefers to keep low inventory of the containers due to their high cost. Therefore, each container has the capacity for 5 bags.

OBJECTIVE: Create a forecasting model to determine the number of containers required for the period of October 1st – October 3rd, 2018.

STEPS TO TAKE:

- **A)** Use historical data from July 1st to September 21st as training data, and September 22nd to September 30th as the testing data to create the forecasting model.
- **B)** Create another model – AR model to forecast the number of containers.
- **C)** Compare the models created in A) and B) based on Root Mean Square Error and Median Absolute Deviation measures.

- **A)** Plots of the bags used for Whole Blood, Platelets, and Packed Red Cells along with the forecasted simple moving average models are shown:

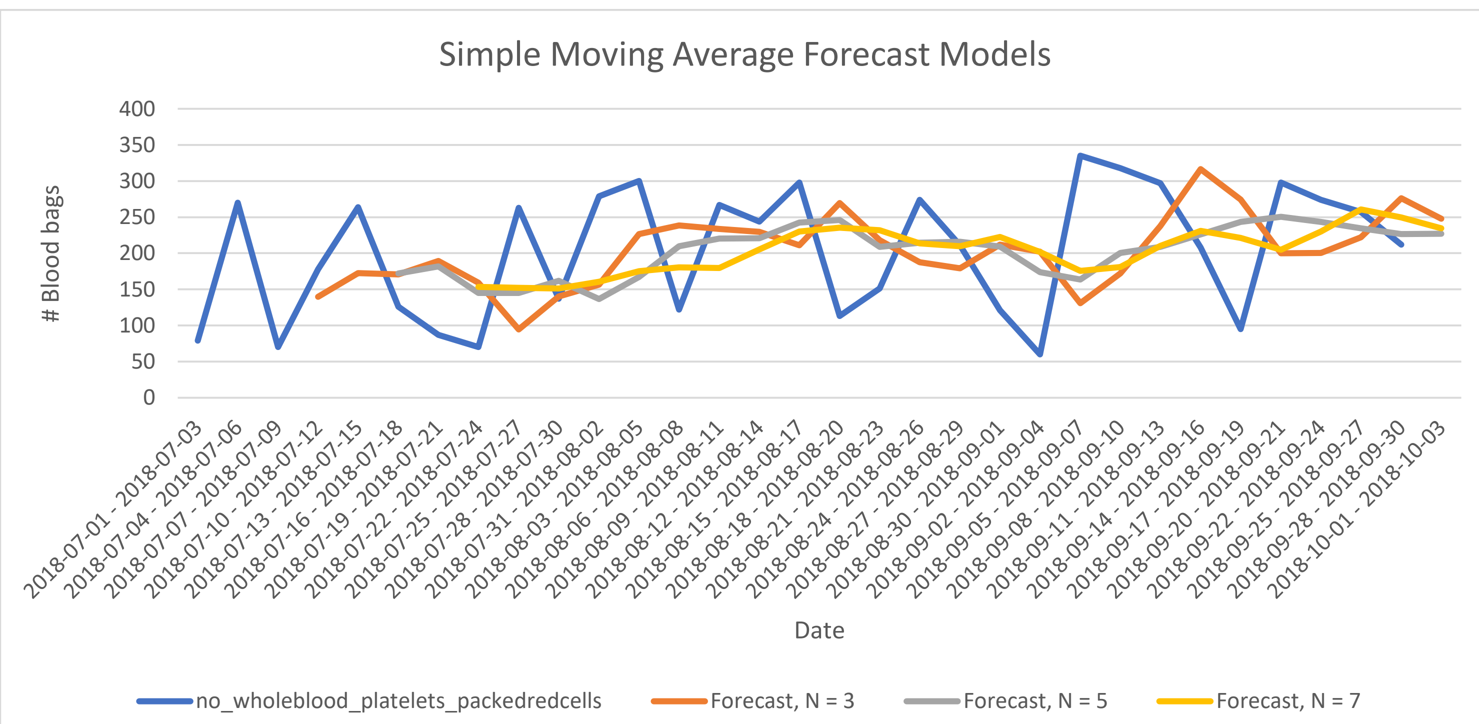

Figure 1. 3-days basis of number of blood bags as well as its forecasted models with N = 3, 5, and 7. Forecasted number of blood bags required for October 1st – October 3rd can be seen in this graph. Converted to containers by dividing number of blood bags by 5.

Raw data and how the simple moving average was created using Excel. This can be seen in the uploaded Excel file with worksheet named "A + B + C." The "N" values were chosen arbitrarily and they represent the number of rows used to calculate the simple moving average. Summary of forecasted number of containers required for October 1st – October 3rd shown below:

Table 1. Forecasted number of containers required for October 1st - October 3rd. It's the total number of blood bags within those 3-days divided by 5. Numbers rounded higher because logically speaking it is better to have an extra container.

2018-10-01 - 2018-10-03	**N = 3**	**N = 5**	**N = 7**
Blood Bags	247.6667	227.2	234.4286
Containers	49.53333 = ~50	45.44 = ~46	46.88571 = ~47

B)

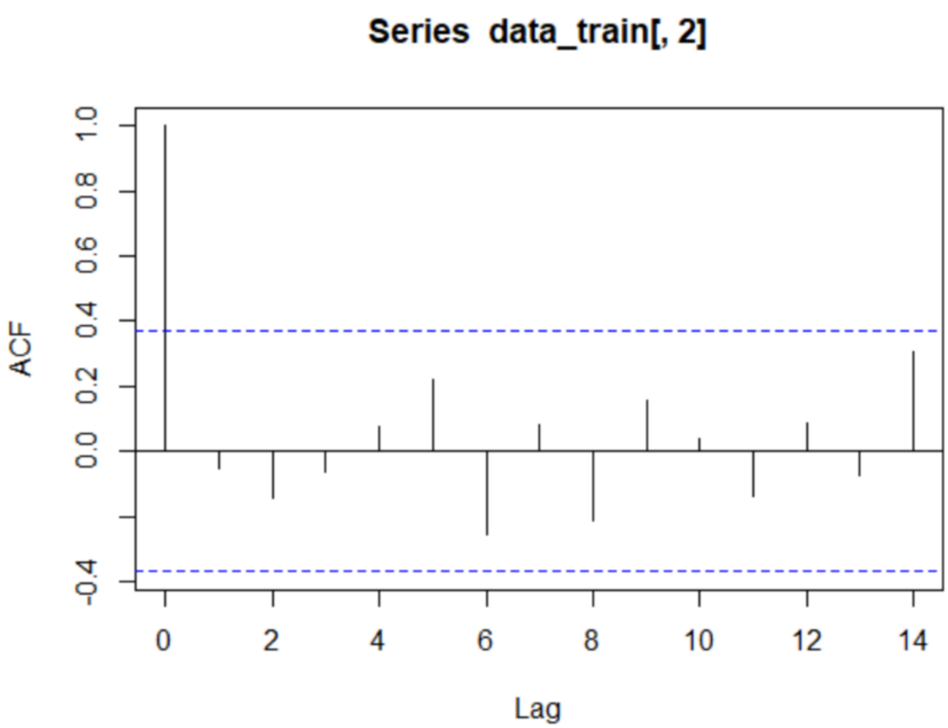

Figure 2. ACF plot concluded that there is no evidence of non-stationarity.

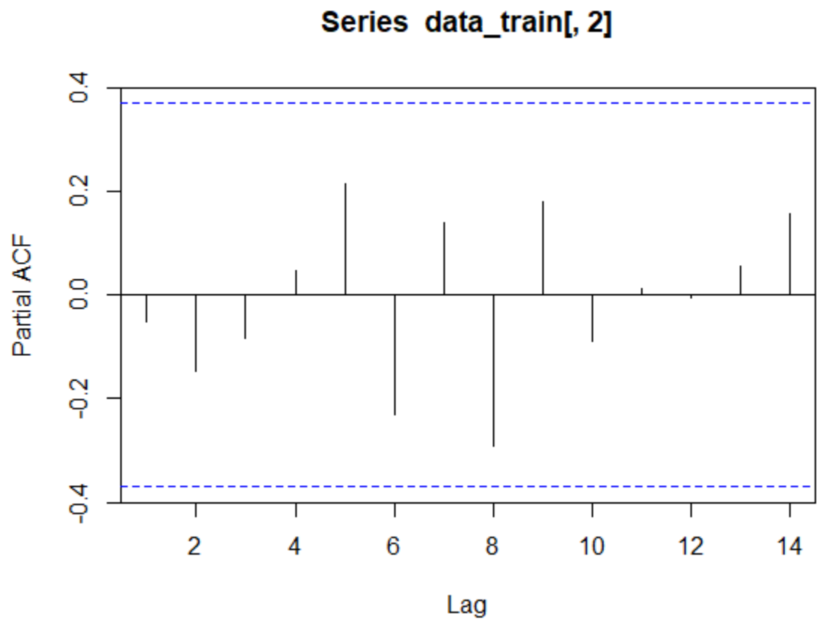

Figure 23. PACF plot shows no autocorrelation of residuals. Therefore, can accept model AR(1).

From the residual ACF and PACF plots as shown in Figure 2 and 3, the AR model with p = 1 can be used. The ACF plot concluded that there is no evidence of non-stationarity, and the PACF plot shows no autocorrelation of residuals, so the model AR(1) can be accepted. Plots were created using R. The AR model was determined to be:

$$\boldsymbol{y_t = 212.1785 - 0.05209 y_{t-1} + \epsilon_t} \text{ or}$$

$$\boldsymbol{y_t = 212.1785 - 0.05209 y_{t-1}}$$

Where y_t is the dependent variable result, *c* is 212.1785, y_{t-1} is the independent variable, and θ_1 is -0.05209.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.052606
R Square	0.002767
Adjusted R Square	-0.03712
Standard Error	91.82046
Observations	27

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	584.9159151	584.9159	0.069377	0.794405104
Residual	25	210774.9359	8430.997		
Total	26	211359.8519			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	212.1785	42.2366057	5.02357	3.51E-05	125.190607	299.1664	125.1906	299.1664
Yt-1	-0.05209	0.197781952	-0.26339	0.794405	-0.459434298	0.355245	-0.45943	0.355245

The model forecasted 201.1344 blood bags, which is equal **to 40.22689 (round up to ~40) containers.**

C) MAD and RMSE measures were determined using Excel and summarized in the following table:

Table 2. RMSE and MAD values determined using Excel for Questions A and B.

	(A) Simple Moving Average			**(B)**
	N = 3	**N = 5**	**N = 7**	**AR Model**
RMSE	59.90919054	23.57089	33.41086	56.71310405
MAD	57.55555556	22.6	28.38095	49.88365586

Based on MAD and RMSE, measures, the simple moving average forecasting model with N = 5 was determined to be the best. This is because the model was determined to have the lowest MAD and RMSE values. The smaller Root Mean Squared Error and Mean Absolute Deviation between forecast and actual results, the better.