**Time Series Analysis of Total US Construction Spending on Manufacturing Sector from 2002 to 2020**
**Aditya Thakur**
**MA 585**


## 1. Introduction

  The functioning of the manufacturing sector is a key indicator of growth for many countries. The workings showcase the level of technological integration that is achieved by a country, with the sector often being the torchbearer for the implementation of automated, cutting-edge technology. A growth in spending on manufacturing not only emphasizes the importance of innovation, but also encourages businesses reliant on the production of goods. Therefore, I have decided to look at how the spending on the construction of manufacturing units in the US has evolved over the past few years.

## 2. Data

  The source of the dataset is the Federal Reserve Economic Data (FRED) database, which contains information on more than 750,000 economic time series from 96 different sources, maintained by the Research Division of the Federal Reserve Bank of St. Louis. The dataset is linked in Appendix 1. After importing the dataset, which contained data from January 2002 to February 2021, I took the following steps to clean the dataset:

  i. Reclassifying columns to fit numerical and datetime formats
  ii. Removing metadata columns containing qualitative information not relevant to this analysis

The first 5 rows of the finalized dataset are shown in Figure 1.


Figure 1: Finalized Dataset (first 5 rows)

| DATE <date> | Amount <dbl> |
|---|---|
| 2002-01-01 | 2134 |
| 2002-02-01 | 1972 |
| 2002-03-01 | 1994 |
| 2002-04-01 | 1971 |
| 2002-05-01 | 1976 |

## 3. Exploration

  A plot of the above data can be seen in Figure 2. As visible, there is an upward trend as time passes, with some drops near 2010 and 2016. There is also a clear seasonal pattern that is visible. One particularly noteworthy observation is the sharp drop in spending near the start of 2020. Much of this decrease can be attributed to the COVID-19 pandemic which has brought multiple economies and workplaces to a standstill.

  On first glance, this dataset does not seem stationary, due to the visible seasonality accompanied by the upward trend. To further investigate this claim, additive decomposition was performed, shown in Figure 3, which confirms the presence of trend and seasonality components. To confirm this visible non-stationarity, I performed the Augmented Dicky-Fuller test, the results of which are shown in Figure 4. As the p-value (0.0842) is greater than

the significance level (0.05), I failed to reject the null hypothesis, which suggested a presence of unit root and, in turn, indicated that the time series is non-stationary.

Figure 2:



Plot of Expenditure on Manufacturing Sector (Construction) from 2002-2021

Figure 3:



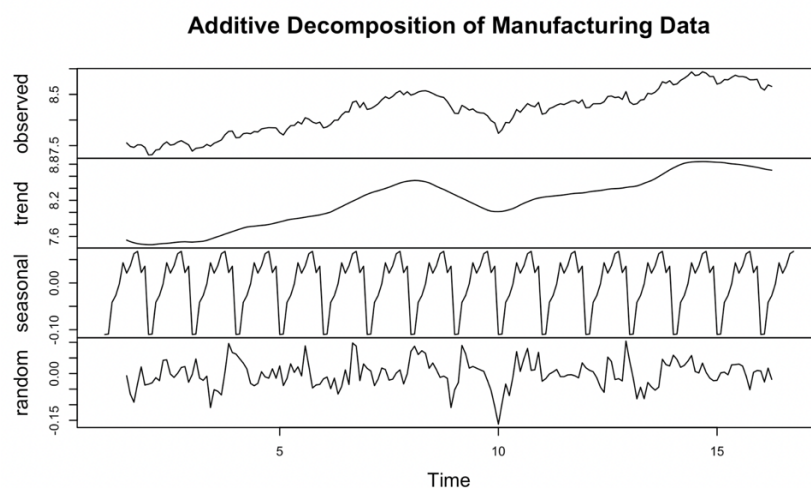Additive Decomposition of Manufacturing Data

Figure 4:



```
        Augmented Dickey-Fuller Test

data:  manu_train
Dickey-Fuller = -3.318, Lag order = 12, p-value = 0.07002
alternative hypothesis: stationary
```

## 4. Transformation

To ensure that the data is stationary, I performed a differencing of order 1 (for trend) and order 12 (for seasonality- since the data is monthly, the period of the data is 12). Figure 5 shows a comparison of the original dataset before and after differencing, and Figure 6 shows the results of the ADF test performed on the differenced dataset. From Figure 5, we can see that the differenced data appears to be centred around 0 (mean is 0)

while the variance is relatively constant throughout the timespan (variance does not depend on time.) This result, combined with the p-value from the revised ADF test (<0.01) allows us to reject the null hypothesis of non-stationarity and allows us to say, with 95% confidence, that this differenced dataset is stationary.
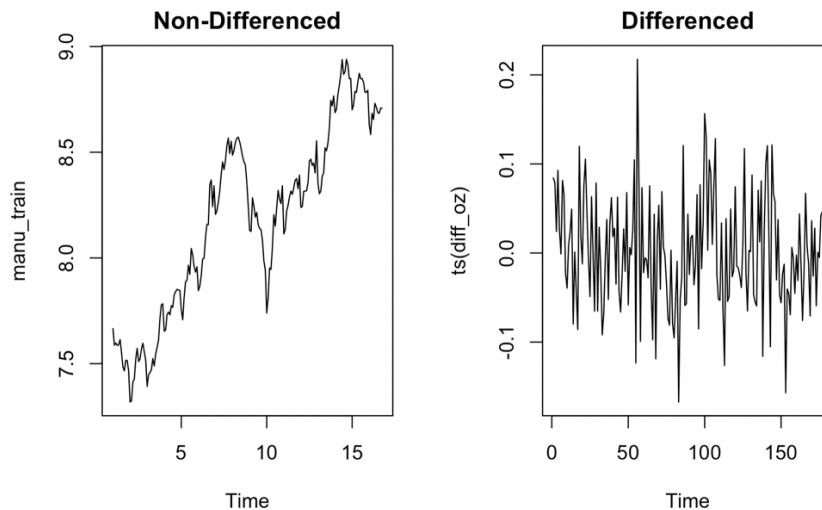
Figure 5:



Figure 6:



## 5. Model Selection

To select an appropriate model, I first observed the autocorrelation and partial auto-correlation function, as shown in Figure 7. From this figure, I first tried to identify an appropriate non-seasonal ARIMA model. The ACF and PACF both cut off after the initial lags, suggesting a low-order non-seasonal ARMA model.

Given that the period of the dataset is 12, we now take a look at the behaviour of the ACF and PACF at the seasonal lags. The ACF is significant at the first 2 seasonal lags (lags 12 and 24) after which it is cut off. The PACF meanwhile is significant at the first 4 lags. This suggests a seasonal MA(2) model.

Due to the uncertainty created by the ACF and PACF plots, it becomes important to compare different models and their AICc values.
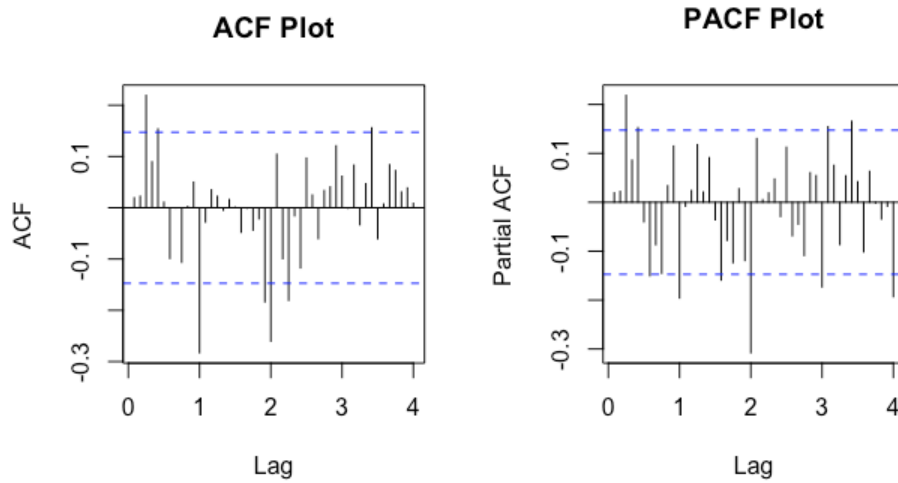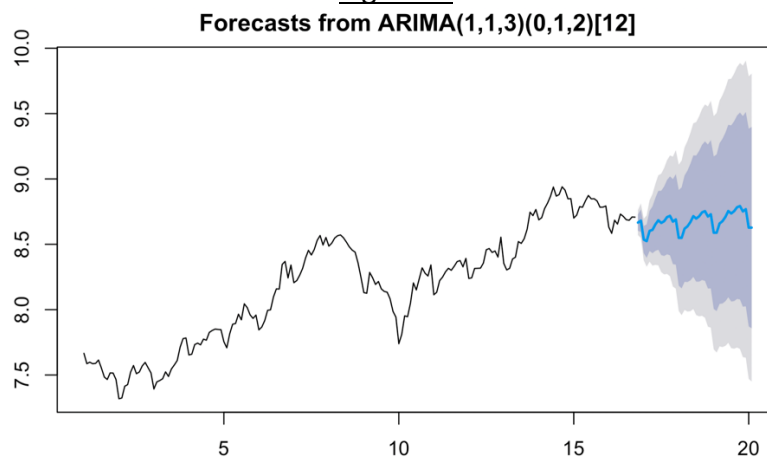
Figure 7:



ACF Plot   PACF Plot

Figure 8 showcases AICc values for different models. After attempting a fit with several low-order non-seasonal ARIMA models, the model with order (1,1,3) has the lowest AICc. Therefore, the final model for this time series is (1,1,3)x(0,1,2)[12]. Figure 9 shows an ARIMA forecast based on this model.

Figure 8:

| Model | AICc |
|---|---|
| (1,1,3)x(0,1,2)[12] | -542.1863 |
| (1,1,2)x(0,1,2)[12] | -534.8231 |
| (1,1,1)x(0,1,2)[12] | -534.5 |
| (1,1,1)x(0,1,1)[12] | -530.662 |
| (1,1,1)x(1,1,1)[12] | -529.26 |

Figure 9:



Forecasts from ARIMA(1,1,3)(0,1,2)[12]

## 6. Model Diagnostics

The estimated parameters for the chosen (1,1,3)x(0,1,2)[12] model are given in Appendix 2. Taking a look at the model diagnostics, shown in Figure 10, we see that the ACF of the residuals is not significant at any lag. Additionally, the p-values are all above the cut-off interval. These findings suggest an adequate model has been fit to this time series.

Additionally, the Q-Q Plot in Figure 11 suggests that the data came from a normal population.
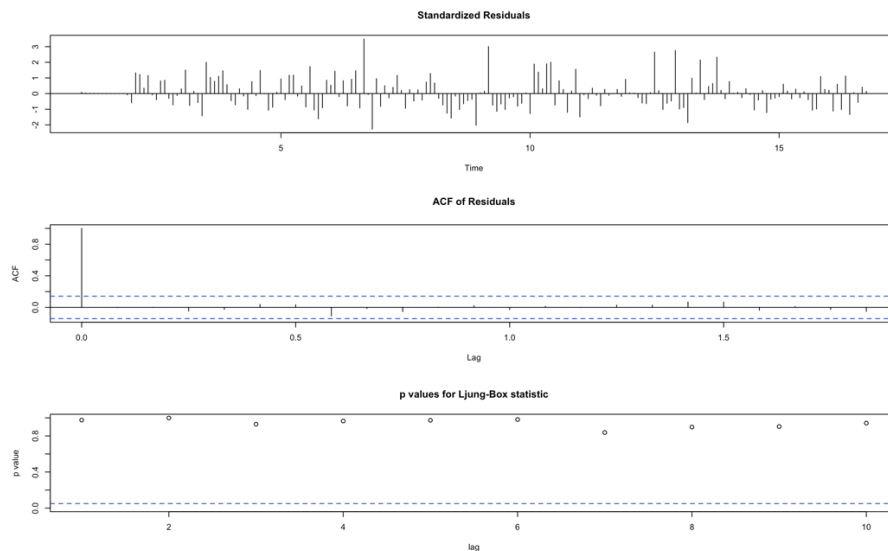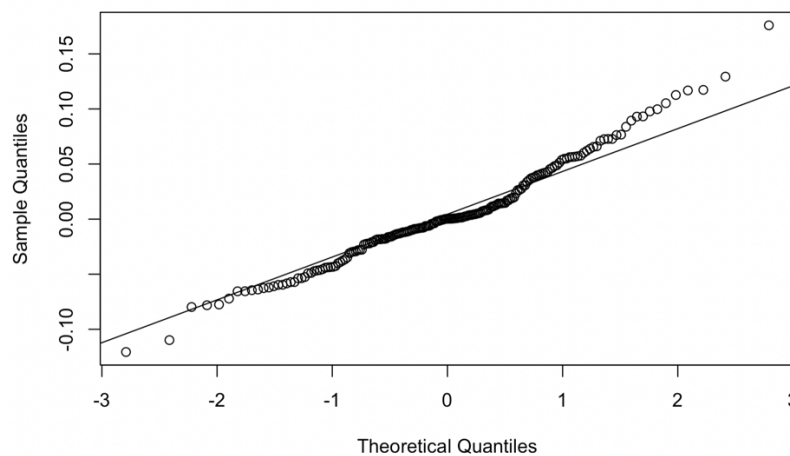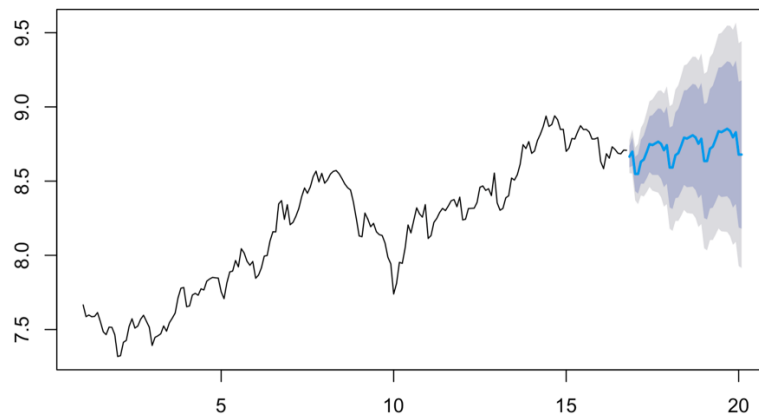
Figure 10:



Figure 11:



## 7. Holt-Winters Forecasting

To provide an alternative method for forecasting future values, I also implemented a Holt-Winters Forecast. The forecast, as shown in Figure 12, was performed using the multiplicative version. This is because the time series showcased seasonality that was increasing as time passed. The forecast estimates are provided in Appendix 3.

The forecast is reasonable, since it captures the seasonal movements well while also recognizing the upward trend that was present near the end of the data. This forecast, much like the ARIMA forecast, does not account for the sudden drop witnessed near the beginning of 2020 due to the COVID-19 pandemic. Given the unprecedented nature of the event, the forecast is expected to have large errors due to this. Nevertheless, the recovering economy has bolstered productivity and the upward trend in manufacturing spending holds true despite the pandemic.

Figure 12:



**Forecasts from HoltWinters**

Comparing both forecasts, the seasonality is well captured in both and there is a slightly more upward trend in the Holt-Winters forecast than the ARIMA forecast. Quantitatively, Figure 13 shows important metrics on the performance of both forecasts.

Figure 13: Forecast Diagnostics

|        | Holt-Winters | ARIMA    |
|--------|--------------|----------|
| **MAE**  | 6160.347     | 6101.085 |
| **RMSE** | 6179.401     | 6119.332 |
| **MAPE** | 99.86        | 99.857   |

After comparing these metrics on the two forecasting methods, it is clear that both forecasting methods have similar accuracy and are comparable.

**8. Conclusion**

From the above analysis, I was able to build a seasonal ARIMA model which I used to forecast future values of construction spending on the manufacturing sector. I conducted ADF tests along with differencing techniques to account for trend and seasonality components and then used the ACF and PACF plots along with AICc values to derive an adequate model for this time series. That model was further used for forecasting.

These analyses and forecast show a general upward trend in the spending on the construction of manufacturing sites. The effect of the COVID-19 pandemic has also been taken into account in this dataset, where a sharp drop can be seen in Figure 2 over the 2020 period. A recovery is being made and forecasted by this model however, due to the unexpected and unprecedented nature of this pandemic, the forecast does not account for the sudden drop in activity that has been witnessed worldwide. A future study, using intervention analysis to account for sudden changes in the manufacturing sector due to COVID-19, is a topic I will be pursuing further.

The upward trend in manufacturing spending is a healthy sign for the US economy. The sector powers various businesses and serves as a benchmark for technological innovation. A swift recovery from the pandemic has boosted integration levels and productivity, restabilising manufacturing as a growing and prevalent sector.

## Appendices:

## Appendix 1: Dataset URL:

https://fred.stlouisfed.org/series/TLMFGCON

## Appendix 2: Parameter Estimates for (1,1,3)x(0,1,2) Model:

```
Series: manu_train
ARIMA(1,1,3)(0,1,2)[12]

Coefficients:
         ar1      ma1     ma2     ma3     sma1     sma2
      0.6231  -0.6666  0.0077  0.2785  -0.7653  -0.2347
s.e.  0.1290   0.1306  0.0852  0.0779   0.1131   0.0889

sigma^2 estimated as 0.002214:  log likelihood=278.42
AIC=-542.85   AICc=-542.19   BIC=-520.62
[1] -542.1863
```

## Appendix 3: Holt-Winters Forecast Values:

| | Point Forecast | Lo 80 | Hi 80 | Lo 95 | Hi 95 |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| Nov 16 | 8.665006 | 8.592493 | 8.737519 | 8.554106 | 8.775905 |
| Dec 16 | 8.698709 | 8.602229 | 8.795190 | 8.551156 | 8.846263 |
| Jan 17 | 8.548317 | 8.432323 | 8.664310 | 8.370919 | 8.725714 |
| Feb 17 | 8.548502 | 8.415469 | 8.681534 | 8.345045 | 8.751958 |
| Mar 17 | 8.630831 | 8.482376 | 8.779287 | 8.403788 | 8.857875 |
| Apr 17 | 8.645576 | 8.482850 | 8.808301 | 8.396708 | 8.894443 |
| May 17 | 8.695115 | 8.518988 | 8.871242 | 8.425753 | 8.964478 |
| Jun 17 | 8.749380 | 8.560533 | 8.938226 | 8.460564 | 9.038196 |
| Jul 17 | 8.743165 | 8.542147 | 8.944182 | 8.435734 | 9.050595 |
| Aug 17 | 8.754398 | 8.541660 | 8.967135 | 8.429044 | 9.079751 |
| Sep 17 | 8.765663 | 8.541585 | 8.989741 | 8.422966 | 9.108360 |
| Oct 17 | 8.751123 | 8.516026 | 8.986220 | 8.391574 | 9.110673 |
| Nov 17 | 8.708481 | 8.459658 | 8.957304 | 8.327938 | 9.089023 |
| Dec 17 | 8.742184 | 8.482978 | 9.001391 | 8.345762 | 9.138607 |
| Jan 18 | 8.591792 | 8.322399 | 8.861185 | 8.179791 | 9.003793 |
| Feb 18 | 8.591977 | 8.312571 | 8.871382 | 8.164662 | 9.019291 |
| Mar 18 | 8.674306 | 8.385042 | 8.963571 | 8.231914 | 9.116699 |
| Apr 18 | 8.689051 | 8.390063 | 8.988038 | 8.231789 | 9.146312 |
| May 18 | 8.738590 | 8.430002 | 9.047178 | 8.266646 | 9.210535 |
| Jun 18 | 8.792855 | 8.474775 | 9.110934 | 8.306394 | 9.279315 |
| Jul 18 | 8.786640 | 8.459166 | 9.114113 | 8.285812 | 9.287467 |
| Aug 18 | 8.797873 | 8.461093 | 9.134652 | 8.282813 | 9.312932 |
| Sep 18 | 8.809138 | 8.463132 | 9.155144 | 8.279968 | 9.338308 |
| Oct 18 | 8.794598 | 8.439437 | 9.149759 | 8.251427 | 9.337770 |
| Nov 18 | 8.751956 | 8.385449 | 9.118463 | 8.191431 | 9.312480 |
| Dec 18 | 8.785659 | 8.410174 | 9.161145 | 8.211404 | 9.359914 |
| Jan 19 | 8.635267 | 8.250853 | 9.019680 | 8.047357 | 9.223176 |
| Feb 19 | 8.635452 | 8.242156 | 9.028748 | 8.033957 | 9.236946 |
| Mar 19 | 8.717781 | 8.315643 | 9.119920 | 8.102764 | 9.332799 |
| Apr 19 | 8.732526 | 8.321582 | 9.143469 | 8.104041 | 9.361010 |
| May 19 | 8.782065 | 8.362349 | 9.201781 | 8.140165 | 9.423966 |
| Jun 19 | 8.836330 | 8.407871 | 9.264788 | 8.181059 | 9.491601 |
| Jul 19 | 8.830114 | 8.392940 | 9.267289 | 8.161513 | 9.498716 |
| Aug 19 | 8.841348 | 8.395481 | 9.287215 | 8.159453 | 9.523242 |
| Sep 19 | 8.852613 | 8.398075 | 9.307151 | 8.157457 | 9.547769 |
| Oct 19 | 8.838073 | 8.374883 | 9.301263 | 8.129685 | 9.546461 |
| Nov 19 | 8.795431 | 8.321679 | 9.269183 | 8.070890 | 9.519972 |
| Dec 19 | 8.829134 | 8.346796 | 9.311473 | 8.091461 | 9.566807 |
| Jan 20 | 8.678742 | 8.187828 | 9.169655 | 7.927954 | 9.429529 |
| Feb 20 | 8.678927 | 8.179447 | 9.178406 | 7.915039 | 9.442814 |