Costco Revenue Forecasting Project Report

San Francisco State University

DS 604

Julie Germenis

Uyen (Ashley) Huynh

Winnie Liu

Erik Wong

May 10, 2021

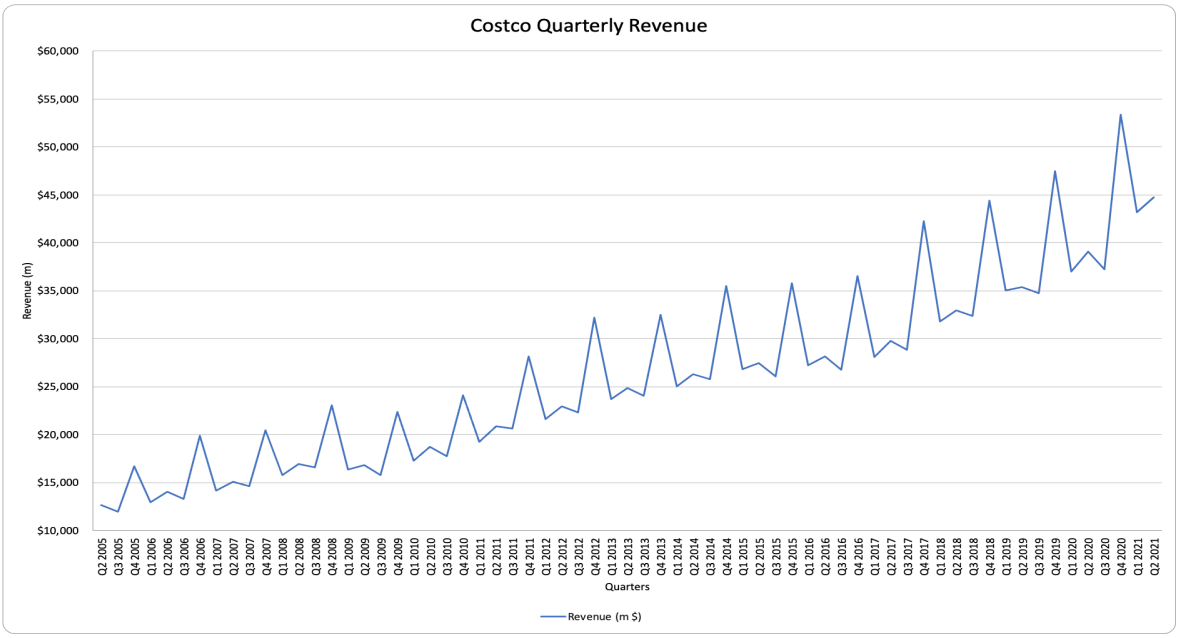
A. Background

A long list of retailers who have reported bankruptcy has made 2020 a year of failure for many corporate businesses in the industry. The list of casualties includes Guitar Center, Pier One, Lord & Taylor, Neiman Marcus, Sur La Table, Brooks Brothers, and Stein Mart. At the same time, consumers increase their reliance on business which is dominantly or partly extended in online shopping such as Amazon, Costco, Walmart, and more. According to eMarketer's forecast, it estimates that the US retail sales are expected to increase 2 percent to $5.574 trillion in 2020, representing a deceleration of the retail market amid rising economic uncertainty (2020). All of these factors are requiring brick-and-mortar businesses to innovate like never before — or risk losing their customers (eMarketer, 2020). Costco Wholesale is a membership warehouse club and a multi-billion dollar global retailer with its club operations in eight countries. Costco provides a wide selection of merchandise to its customers. The company first opened under the name Price Club in 1976 in San Diego. Originally its customer was small businesses, and the company decided to switch its market by serving a selected audience of non-business members. Since 1997, the company has grown worldwide, with total sales in recent fiscal years exceeding $64 billion. This background makes Costco an excellent candidate for our project of forecasting how well a retailer company performs in the pandemic.

We obtained the Costco data through the Macrotrends website [1]. The data set projects the fiscal year with data from 2005 through 2021. We recognized that the data is seasonal, with the highest quarter in quarter 4. There is a trend of an increase in Costco revenue during this quarter. It appears that Costco's revenue has been growing during the Covid-pandemic, but this does not affect the seasonal data.[1]

---

[1] *Costco Revenue 2006-2021: COST*. Macrotrends. (n.d.). https://www.macrotrends.net/stocks/charts/COST/costco/revenue.



B. A description of the models and tools

1. Winters' Model

We used Excel to run Winters' exponential smoothing and Decomposition method by looking at the MAPE value to see which way can perform a better forecast for 2021 Q2 to 2022 Q2 Costco sales.

There will be three smoothing constants in the Winters' method: alpha for the level, gamma for the trend, and beta for the seasons.

To initialize:

We used the naive method to set the initial value of level (Ft), trend (Tt), and seasonal indices (St), starting from Q1 2006. We took actual revenue Q1 2006 for the level initialization and 0 for the trend initialization. Then, we used a more sophisticated initialization for seasonal indices due to the high seasonality that Costco has shown in the data, with Q4 being significantly higher in sales than the other quarters. Thus, our seasonal indices initialization will be the actual sales divided by the average of four quarterly sales in 2005. As we can see, the fourth quarter's

seasonal index is 1.23, which means its revenue is 23 percent higher than the average quarterly revenue.

Costco Quarterly Revenue
(Millions of US $)

Date	Quarter	Revenue (m)	Ft	Tt	St	Wt
2/28/05	Q2 2005	$12,658			=C4/AVERAGE($C$4:$C$7)	
5/31/05	Q3 2005	$11,997			0.884	
8/31/05	Q4 2005	$16,719			1.231	
11/30/05	Q1 2006	$12,933	$12,933	0	0.953	
2/28/06	Q2 2006	$14,059	$14,473.38	$165.64	0.950	$12,057.81

Then, we set up some random values of alpha, gamma, and beta between 0 and 1. Next, we will run the solver to get the optimal solution. Finally, we are able to apply the actual sales and the parameters into equations for level, trend, seasonal index, and Winters'. Since we were going to determine our final model based on the value of MAPE, we needed to take the difference of actual sales and the value of Winters' and divide by actual sales in absolute value.

Costco Quarterly Revenue
(Millions of US $)

Date	Quarter	Revenue (m)	Ft	Tt	St	Wt	absolute % error
2/28/05	Q2 2005	$12,658			0.937		
5/31/05	Q3 2005	$11,997			0.884		
8/31/05	Q4 2005	$16,719			1.231		
11/30/05	Q1 2006	$12,933	$12,933	0	0.953		
2/28/06	Q2 2006	$14,059	$14,473.38	$165.64	=$K$3*(C8/D8)+(1-$K$3)*F4		14.234%
5/31/06	Q3 2006	$13,284	$14,921.92	$196.06	0.887	$12,935.67	2.622%

Alpha	0.72
Gamma	0.11
Beta	0.46
MAPE	2.63%

In this Winters' model, the lowest MAPE we could get after we ran the solver is 2.63 with a relatively high alpha value of 0.7. The two-way data table shows the optimal combination of alpha and gamma, which run at 0.7 and 0.1.

Alpha	0.72
Gamma	0.11
Beta	0.46
MAPE	2.63%

Data Table									
alpha across, gamma down									
2.63%	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	5.14%	3.90%	3.28%	2.99%	2.78%	2.68%	2.64%	2.66%	2.89%
0.2	4.58%	3.55%	3.12%	2.88%	2.75%	2.68%	2.64%	2.70%	2.98%
0.3	4.42%	3.41%	3.09%	2.90%	2.81%	2.70%	2.69%	2.79%	3.12%
0.4	4.28%	3.40%	3.11%	2.98%	2.88%	2.75%	2.76%	2.91%	3.28%
0.5	4.14%	3.44%	3.18%	3.05%	2.94%	2.80%	2.88%	3.07%	3.45%
0.6	4.06%	3.47%	3.24%	3.11%	2.99%	2.87%	2.99%	3.23%	3.61%
0.7	4.00%	3.49%	3.31%	3.18%	3.05%	2.95%	3.10%	3.40%	3.79%
0.8	4.01%	3.54%	3.37%	3.23%	3.13%	3.03%	3.21%	3.62%	4.00%
0.9	4.04%	3.59%	3.42%	3.34%	3.21%	3.12%	3.34%	3.91%	4.22%

Considering the impact of the pandemic, we were not sure whether the year of 2020 would be an outlier. Hence, we calculated the MAPE without the year of 2019, 20, and 21. And the new in-sample MAPE is 2.7 percent, where the out-sample MAPE, which contains the year of 2019 to 2021 Q1 is 2.2 percent (Calculation is in "Winters (2)" Sheet). The result tells us that the sales of 2020 is realistic and valid for us to include the numbers of 2020 for forecasting

mape (in-sample)=	2.70%	2005-2018
mape (out-sample) =	2.212%	2019-2021

future quarterly sales.

To Forecast:

We took the last quarter of data to forecast 2021 Q3, Q4, 2022 Q1, and Q2. The forecasted revenue of Q4 is still higher than other quarters, which makes sense.

Quarters	m	Ft	Tt	St	Wt
Q3 2020				0.912	
Q4 2020				1.246	
Q1 2021				0.944	
Q2 2021		$46,566.58	$997.25	0.960	$44,427.44
Q3 2021	1				$43,357.20
Q4 2021	2				$60,526.69
Q1 2022	3				$46,799.99
Q2 2022	4				$48,547.66

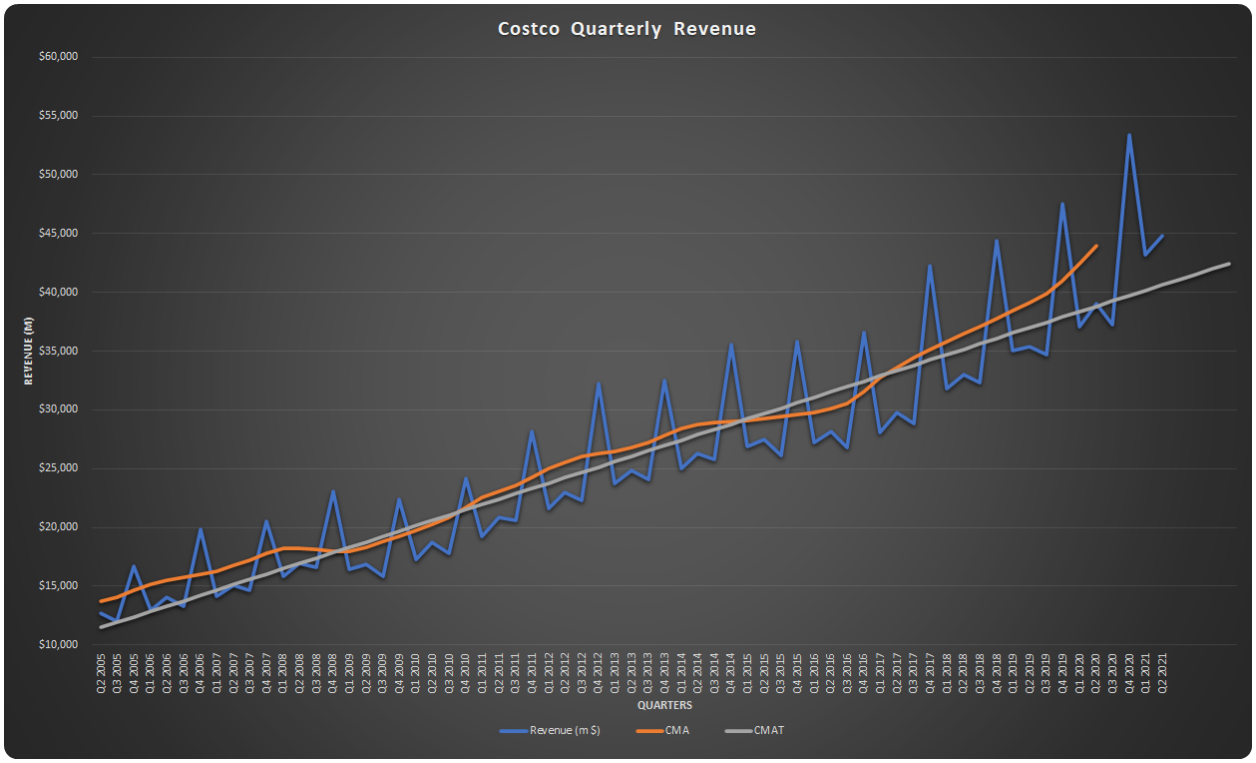
2. Decomposition Model

Using Excel, we built the second model, the decomposition model, before deciding on our final model. Since we have seasonal data, we wanted to isolate and examine these data patterns individually. The data shows that the fourth quarter is a high-revenue quarter, while the third quarter is a low-revenue quarter.

First, we found the moving average by averaging revenues of 4 quarters respectively in order. Then, we calculated the center moving average by averaging two of the moving averages respectively in order.

Revenue (m)	MA	CMA
$12,658		
$11,997		
$16,719	$13,577	$13,752
$12,933	$13,927	$14,088

After finishing the model, we have the deseasonalized data represented by the centered moving average. The first moving average helps to remove any irregular pattern in the data. The centered moving average helps to smooth out any random fluctuation in the data. As you can see in the graph below, the Costco revenue has an upward trend represented by deseasonalized data.



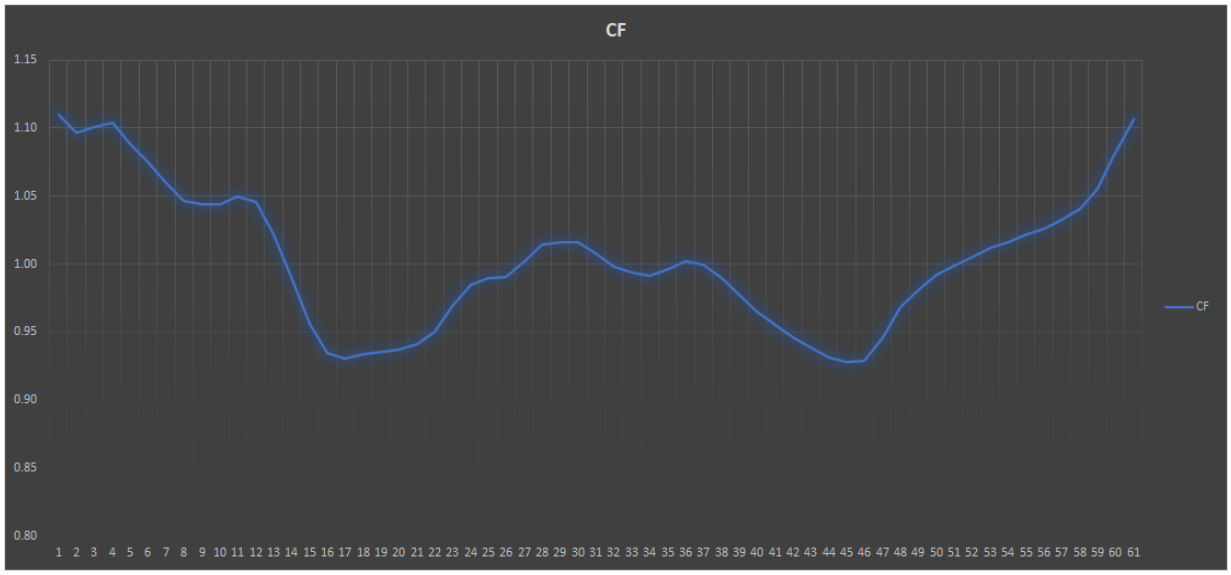
After that, a long-term trend is calculated for the deseasonalized data by finding a linear equation, CMAT, and by running a regression with time index as an X variable and CMA as a Y variable. The result is:

CMAT = 11025.21 + 455 * (Time Index)

Where: time index = 1 for the first quarter and increased by 1 each quarter thereafter.

The long-term trend line, CMAT, shows a growing trend in Costco's revenue.

We also know that a cycle factor greater than 1 indicates that the deseasonalized value for the given period is above the long-term trend of the data and vice versa. The historic cycle factor of the Costco quarterly revenue is shown in the CF graph below. From the result, approximately 52% of the data is under or equal to 1, while 48% of the data is greater than 1. It indicates that about half of the deseasonalized data is above the long-term trend.



Since we can only calculate the cycle factor until Q4 2020, we assigned numbers from Q1 2021 to Q2 2022 to help the forecasting process. We decided on the number by reasoning from the current issue and the historical data. We saw an increasing trend in the cycle factor since Q1 2020, and it kept increasing until Q4 2020; thus, we predict the trend will continue to rise until Q1 2021. However, we restricted the numbers of cycle factors between 1.09-1.11 for the rest of the quarters. We determined based on the highest period data from Q4 2005 to Q4 2006.

The calculated MAPE for this method is 1.20 percent lower than 2.63percent from the Winters' model.

We also take the pandemic into account as a factor that will make the sales in 2020 an outlier in this model, so we calculated in-sample MAPE (without the year 2019, 20, and 21) and out-sample MAPE (which includes the year 2019, 20 and 21). The result is shown in the table below. We also take the pandemic into account as a factor that might make the sales in 2020 an outlier in this model, so we calculated in-sample MAPE (without the year 2019, 20, and 21) and out-sample MAPE (which include the year 2019, 20 and 21). The result is in the table below (Calculation is in “Decomposition (2)” Sheet).

Mape (in-sample)	**1.09%**	**2005-2018**
Mape (out-sampe)	**1.87%**	**2019-2021**

The in-sample MAPE of 1.09 percent is lower than the out sample MAPE of 1.87 percent. The difference, 0.78 percent, between these two MAPE is lower than 1 percent, which indicates that the revenue of 2020 is realistic and reasonable to be included for forecasting future quarterly sales. The forecast value is denoted based on the model as Forecast. Thus,

$$\text{Forecast} = (\text{CMAT})(\text{SI})(\text{CF})(\text{I})$$

Where I is assumed equal to 1, since we have reason to believe the 2020 data is not an outlier to the model. The forecasting for Q4 2021 is still higher than other quarters, which shows a constant pattern with the actual data.

Date	**Quarter**	**Month**	**Time Index**	**Forecast**
2/28/2021	Q2 2021	Feb	65	**$42,812.89**
5/31/2021	Q3 2021	May	66	**$40,540.78**

8/31/2021	**Q4 2021**	**Aug**	**67**	**$56,185.05**
11/30/2021	Q1 2022	Nov	68	**$42,497.07**
2/28/2022	Q2 2022	Feb	69	**$43,926.55**

C. A discussion of the results

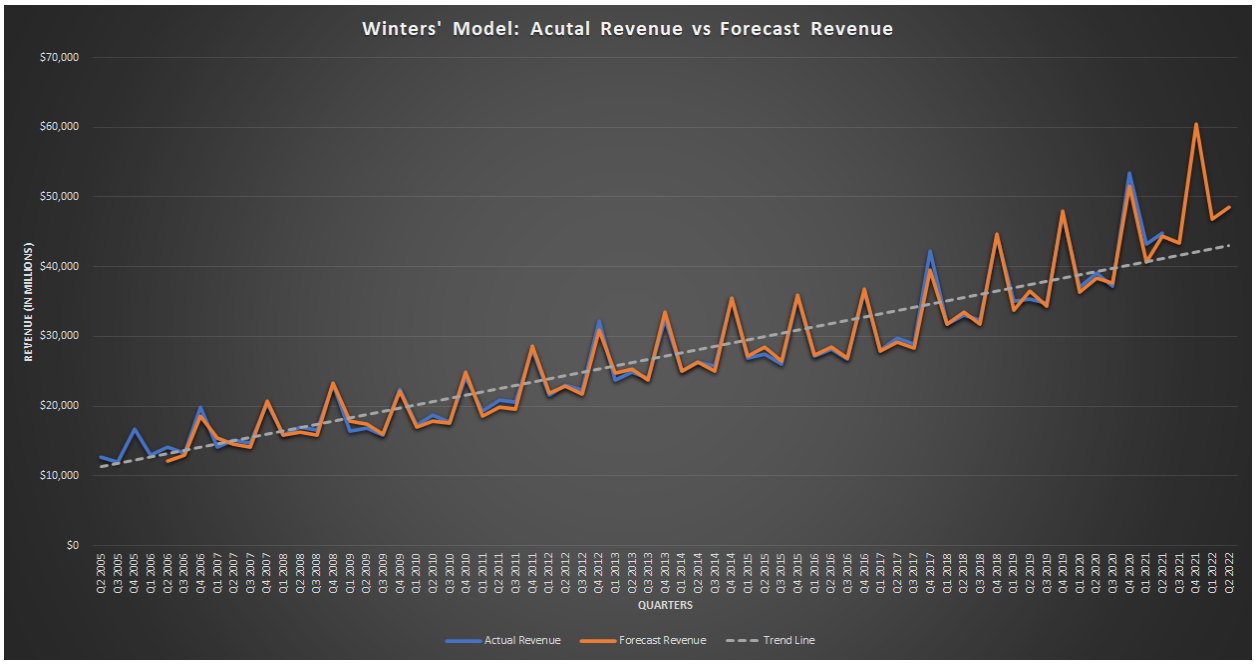
In this work, we have time series and seasonal data. The time series is determined quarterly, with the first-quarter ending in November, and the fourth-quarter ending in August. The seasonality shows the highest quarter is the fourth quarter, while the lowest quarter is the third quarter. After running two models, the Winters' and Decomposition model, the results are reported in the table below.

Model	MAPE	MAPE (in and out-sample)
Winters'	2.63%	In-sample MAPE: 2.70% Out-sample MAPE: 2.212%
Decomposition	1.20%	In-sample MAPE: 1.09% Out-sample MAPE: 1.87%

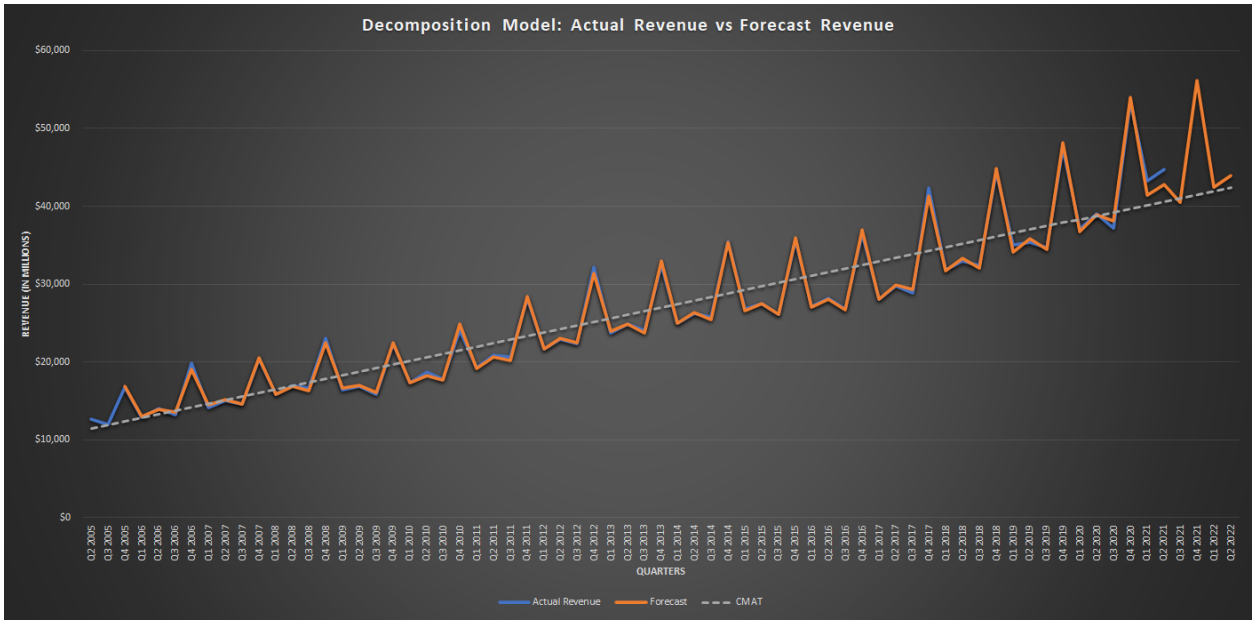
Note that we use MAPE, Mean Absolute Percentage Error, as our measure of accuracy. The smaller the value, the better the fit of the model.

For the Winters' method, the historic MAPE is 2.63 percent, indicating a higher error in the historical period than the decomposition method of 1.20 percent. The in-sample MAPE of 2.70 percent is slightly higher than the historic MAPE, while the out-sample MAPE of 2.212 percent is lower than both of them. The out-sample MAPE indicates that the model is doing pretty well in predicting from the eight quarters from 2019 to the first quarter of 2021. The "Winters' Model: Actual Revenue vs. Forecast Revenue" graph shows that the Winters' method

delivers a moderate forecast for the holdout quarters. The forecast line is slightly under or above the actual line for the holdout period. The latest forecast (from Q3 2021 to Q2 2022) is above the trend line.



We can see that the historic MAPE is 1.20 percent for the decomposition model, indicating a lower error in the historic period. The in-sample MAPE of 1.09 percent is lower than the historic MAPE, which tells us that the model predicts values in the sample effectively within the historical data set. The out-sample of 1.87 percent, on the other hand, is slightly higher than the historic MAPE. However, the fit and accuracy in the decomposition method are constant, which is lower than 2 percent, suggesting that this model is doing quite well. In the "Decomposition Model: Actual Revenue vs. Forecast Revenue", the decomposition method provides a reasonable forecast for the eight quarters from 2019 to the first quarter of 2021. The forecast data is more fitted in with the actual data for the holdout period. The latest forecast (from Q3 2021 to Q2 2022) is much close to the trend line.



The result reveals that MAPE for decomposition is less than that of the Winters' method from this discussion. It also shows that, for the data given, the Decomposition method outperformed the Winters' method.

D. Comments on implementation issues and savings to the company

With our recent finding using the Winter and Decomposition method, we can conclude that Costco will have another stellar year. With quarter four being their peak season. We also found that using a decomposition method will be the best method to use to forecast the Costco season out with a 1.2 percent error. We suggest Costco keep the same system they have been using to run their company. We only urge caution to Costco because of the unpredictable year 2020 has been. We are unsure if that will affect the revenue around for the 2021 and 2022 fourth-quarter forecast.

With the decomposition method as our choice to perform for this data set, we should consider some of the challenges we might experience in the future. As we have known, the smoothing process has a cost in terms of losing some data points. Since a four-period moving

average is used, two points will be lost at each end of the data series by the time the center moving averages have been calculated. Thus, there is no estimate for those quarters for the same period. We also see that the deseasonalized data tends to smooth the rapid rises and falls in the data. In this model, there is an assumption that the seasonal component repeats from year to year. The irregular component is assumed equal to 1 as we have no reason to believe an outlier may take place. This assumption might be reasonable for this given data set. However, we should be more cautious in a longer-series data set. For example, Costco has been growing in revenue, while the global economy went through a downturn due to the Covid-19 pandemic. This trend contrasts with our expectation to see a decrease in Costco's revenue. This result could be different if there were another pandemic that would negatively affect Costco's business operations. The deseasonalized data might not be able to capture these unexpected changes over time.

E. Other comment and suggestion

In the two models we have run, Winter's and Decomposition, we calculated in-sample MAPE and out-sample MAPE to determine that 2020 revenue is a fit to run on these models. In addition, it is assumed to have no irregular factor due to its random nature. We would recommend running the model without the data of 2020 4-quarters and the first quarter of 2021. In other words, this assumes that 2020 data is an exception that needs to be set aside and to carefully examine the data set. After finishing the models, we can better perceive the accuracy and fit of the new model compared to the two models we have built.

Works Cited

eMarketer. (2020, February 24). *The future of retail in the US: industry trends and market trends*. Business Insider. https://www.businessinsider.com/future-of-retail-market-trends.

Costco Revenue 2006-2021: COST. Macrotrends. (n.d.). https://www.macrotrends.net/stocks/charts/COST/costco/revenue.