

Market Potential and Sales Forecasting

There's an old saying derived from a Danish proverb that goes, "It's difficult to make predictions, especially about the future." As difficult as predicting the future is, it's common in business and especially important to marketing. Because marketing is the part of business primarily responsible for generating revenue, forecasting the success of marketing activities has implications for virtually every other part of the business. Sales forecasts affect hiring, investments, salaries, purchasing, production, and just about anything else a business does. Accurate forecasting gives businesses a distinct advantage over competitors who do not prepare forecasts or those who prepare them poorly. Accurate forecasts help businesses better allocate, and hopefully, earn more from their resources.

Apart from the difficulties of predicting the future generally, forecasting itself often evokes nervous reactions because the techniques can be very complex and mathematical. However, not all forecasting need to be so. Many forecasting methods rely solely on informed opinions, while some utilize relatively quantitative analyses. In these notes, we'll overview several different approaches to forecasting.

MARKET AND SALES POTENTIAL

We begin our discussion with a look at potential estimation. While technically not a forecast, potential is closely related to sales forecasting. While forecasting addresses the question "How much will we sell?" potential estimation asks "How much could be sold?" There are two types of potential estimates, though we will ultimately focus on only one. The first is called *market potential*, which will be our focus for most of this section. Recall that a market refers to the people with a want or need for the benefits of a particular product category and the ability to satisfy that want or need. Market potential is defined as the total amount of all brands in a product category that could possibly be sold to the market. The second type of potential estimate is called *sales potential*, which captures the same basic idea as market potential but as it applies to a single brand. Sales potential is defined as the total amount of a single brand that could possibly be sold to the market.

When faced with these definitions, some people may ask, "Why doesn't market potential equal sales potential?" After all, isn't the amount that a single brand could possibly sell actually equal to the amount that could possibly be bought if the brand put all others out of business? Well, in theory, the answer to the question could be yes. However, in reality, probably not. For example, suppose you were tasked with estimating the sales potential of Brand A. People in the product market who were highly loyal to Brand B would be part of the market potential estimate but not Brand A's sales potential estimate. Likewise, it could be that Brand A is not distributed in all of the places that Brand B is distributed. People in the product market with access to Brand B would be part of the market potential

estimate but not part of Brand A's sales potential estimate. Brand A may also lack the product capacity to serve the entire product market. Thus, in practice, there is almost no probability that in competitive product markets the estimates for sales potential will equal the market potential estimates. Forecasts, which will be discussed later, are the predictions of how much will actually be sold during a given time period. The market forecast is the prediction of how much of all brands in a product category will be sold in a given time, while sales forecasts predict sales of a single brand.

For the remainder of this section, we will look at several means of calculating market potential. Our focus will be on market potential and not sales potential for two reasons. Most sales potential estimates are derived from market potential estimates. That is, to calculate sales potential, you generally must calculate the market potential first. Sales potential is typically expressed as a percentage of market potential based on market share predictions. Second, in practice there is very little difference between sales potential and actual sales forecasts. Generally companies try to capture as many customers as they realistically believe they can. If so, then their sales forecast will essentially equal the sales potential estimate.

Market potential is frequently used to estimate whether or not expansion into new markets will be feasible. Markets in this case are most frequently defined geographically, but need not be so. For example, a company may be considering expansion into parts of the country where their product is not currently sold. Therefore, they may want to estimate the potential to sell to customers in these new locations. However, market potential can also be estimated when a firm wants to sell to new demographic groups. For example, the NFL recently began offering team jerseys cut and styled for women because research indicated that many women fans would rather not wear team apparel styled for men. This represents expansion into a new demographic market. Demographic and geographic market definitions are commonly applied together when estimating market potential.

Chain Ratio Method of Calculating Market Potential

The chain ratio method kind of assumes that what happens for one market will happen proportionally in another market. So if we have information on sales or market share in one market, we can extrapolate those results to another market. One common application of the chain ratio method is where a local or regional company wants to calculate potential sales nationally. In this case, the chain ratio method would follow several steps.

First, identify the demographic characteristics of the customer groups to whom the analysis applies. Because the chain ratio method relies heavily on secondary data, often population data from the Census Bureau, it's best to describe markets in terms of sex, age, and income. Although we may use more complex demographic descriptions for other purposes, data for the chain ratio method will be more available, timely, and accurate if the demographic description is kept simple. Sometimes data giving estimates of population size by occupation and some psychographic variables are available. If the data are available and reliable, then they can be used. If not, keep the market description simple.

For example suppose a brand of expensive upscale cookware produced in California sells primarily to women between the ages of thirty-five and sixty-four with household incomes

above \$75,000 annually. This demographic group is simply defined and does not include extra variables such as marital status, family size, education, or occupation. Suppose the company sells through retailers up and down the West Coast.

Second, estimate how many current customers your brand has in the target demographic group. If the group is relatively new, the estimate may be more tenuous. If your company is seeking to sell to the same demographic group but in another area, then the information ought to be easy find internally.

For example, suppose the upscale cookware company had ten thousand purchases by people in that demographic group. How does the company get access to that kind of information? Well, it could survey its dealers or distributors if it uses traditional retail channels. If it uses primarily online distribution, it could survey purchasers as they buy. Or it could rely on syndicated survey or panel data provided by a commercial marketing research provider. No matter what the source, every company with even a modestly sophisticated understanding of its customers ought to have access to information that describes its customers.

Third, divide total sales of your brand to members of the target demographic and then divide by the number of customers in the target demographic to obtain the sales per customer.

For example suppose that the company estimated that it sold five million dollars' worth of cookware to the target demographic group. Assuming that the ten thousand purchases were made by separate individuals, the company would estimate the average purchase per customer to be \$500 ($5,000,000 \div 10,000$).

Fourth, estimate the number of people in the target group that live in the area to which your brand plans to expand and then multiply that number of people by the average sales per customer in that group calculated in the previous step.

For example, suppose the cookware company wants to expand into Arizona, Utah, and Colorado. Using Census Bureau data, the company learns that in these three states there are approximately 406,000 women between thirty-five and sixty-four living in households with annual incomes above \$75,000. The market potential calculation is very straightforward. Multiply the population of the target audience in the new location by the amount spent per year by current customers of the same demographic description. Thus, $406,000 \times 500 = \$203,000,000$.

To emphasize again, this figure is not the sales forecast; it is not a prediction of what they will sell. The market potential estimate is the amount that could possibly be sold if all members of the target audience decided to buy.

The main advantage of the chain ratio method is its simplicity. The data are usually readily available and the mathematics are not complicated at all. The biggest drawback to using the chain ratio method is its main underlying assumptions, which is that the same average purchase rate will hold from one area to another. Of course, companies have little way of knowing that. However, having an empirically based

estimate as a starting point for market expansion planning is better than simply making those decisions from intuition alone.

Buying Power Index

We begin with the simplest method for estimating market potential, excluding just guessing, of course. The Buying Power Index or BPI was initially developed by *Sales and Marketing Management* magazine. The index is very easily calculated with readily available data and is intended to help marketers compare the retail purchasing power of specific locations in the United States. The BPI is not specific to any product or product category. It simply gives a comparative measure of consumer purchasing power relative to purchasing power nationwide. To understand the BPI, let's begin with its formula.

$$BPI = 0.2(\% \text{ of U.S. Population in Area}) + 0.3(\% \text{ of U.S. Retail Sales in Area}) + 0.5(\% \text{ of U.S. Disposable Income in Area})$$

You can see by looking at the formula that it is a weighted sum that describes an area's purchasing power relative to the United States. To use as a measure of market potential, BPIs can be calculated for several areas of similar size or population and then they can be compared to each other. Note again that BPI is not specific to any particular product; it describes a given geographic region. Data to calculate the BPI are readily available from the Census Bureau and from the Bureau of Economic Analysis. BPI figures are most frequently calculated for metropolitan areas generally at the county level.

Regression Based Index Methods

The BPI described above is a very simple method for figuring market potential. As such, it suffers from limitations, not the least of which is its breadth. It is not specific to any product category, but refers to retail spending in general. With a little research, often through internal and external secondary data, marketers can uncover data that provide market potential insights that are specific to their product categories. Of course, the issue here is whether the extra effort to obtain and analyze the data produce results that are sufficiently explanatory to be worth the effort. Like the BPI, regression based index methods convert variables to percentages of the relevant population with the given characteristic, hence the term "index." Unlike the BPI, regression based index methods are not "one size fits all," but are researched and customized to fit the given situation.

In general, regression based index methods follow several steps. First, researchers should consider what factors may predict or explain demand for the product. This requires understanding markets well, which of course is not always easy. Researchers should in particular focus on population factors related to product sales.

For example, suppose a company is considering launching a low cost wireless internet service for families. Pricing may be based on the number of people in the household rather than on data usage. Because the company must install equipment to provide the service, they wish to focus on states with higher population densities so that more people can be reached with less equipment investment. However, it could be that more densely populated areas have more

public spaces with internet, so this variable is in question. Ultimately, they decide to see whether population density, family size, head of household education, and household income predict current internet service sales.

Second, the researchers must collect the data and then create indexes by converting the data to reflect the percent of the base population that possess the characteristics under study.

For example, the researcher studying the internet service would need to obtain data by state on the following variables: sales of internet provider services by state (which would probably be available from the industry association or perhaps even from state regulators), population density, family size, education of the reported head of household, and household income (all of which are available from the Census Bureau). Convert the data to population base percentages would be relatively straight forward. In this case, the base population would be the U.S. population, so for each variable, the researcher would need to divide the state figure by the U.S. population figure. For internet service sales, sales in each state would be divided by nationwide sales to yield the percentage of sales in each state. Similar index calculations would be made for the other variables to indicate the degree to which the state was above or below the average for the U.S.

Third, the researcher would enter the data into a regression analysis and calculate which variables predict internet service sales and the nature of the relationships.

For example, using internet service sales as dependent variable and population density, household education, and income, and family size as independent variables, the researcher would estimate the regression equation to see which independent variables were statistically significant. Independent variables that were not significant could be removed from the model and the equation re-estimated. The SPSS output below shows the regression coefficients from such an estimation. The first estimation showed that household size was nonsignificant. That variable was excluded and the model was re-estimated. Output from the second estimation is below. The estimation yielded an R^2 of 73.9%, suggesting good overall model performance.

Coefficients ^a					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	.478	.077		6.213	.000
1 Density index	.018	.004	.358	4.153	.000
Education index	.266	.086	.297	3.106	.003
Income index	.239	.051	.425	4.718	.000

The equation estimated by the model would be as follows:

$$\text{Sales Index} = .478 + (.018 \times \text{Density}) + (.266 \times \text{Education}) + (.239 \times \text{Income})$$

Note that the independent variables are the indexes of population density, education, and income. These are the variables expressed as a percent of the total population's mean of that characteristic.

Fourth, to compare the market potential for particular states, the data for the states would be put into the equation to calculate the sales index for that state.

For example, suppose that the internet service company wants to compare the market potential indexes for Massachusetts and Rhode Island. The calculations would yield the following:

$$\text{RI: Sales Index} = .478 + (.018 \times 2.79) + (.266 \times 1.06) + (.239 \times 1.38) = 1.14$$

$$\text{MA: Sales Index} = .478 + (.018 \times 2.02) + (.266 \times 2.59) + (.239 \times 1.41) = 1.54$$

The calculations suggest that Massachusetts would be a somewhat better market for the new internet service than Rhode Island. Importantly, this method of market potential estimation does not give an actual estimate of market potential. It is an index that helps compare locations for their potential relative to the average nationally. In this respect it is similar to the BPI. However, in this case, the estimates are specific to the industry and variables that matter in terms of purchase. Therefore, while it is somewhat more difficult to use than the BPI, it does give results that are likely more relevant to marketing managers.

FORECASTING

Forecasting is more fraught with uncertainty because of its fundamental difference from market potential estimation. Forecasting focuses on how much people will buy and not on who could buy. Therefore, as noted earlier, forecasting is about making predictions of future behavior. Forecasting methods range from very simple guesses to highly complex econometric models. In this section, we will focus on some of the simpler methods of forecasting.

Judgmental Forecasting Techniques

As the name implies, judgmental forecasting techniques rely on the judgment of knowledgeable people to make informed guesses about sales in the coming period or periods. Judgmental methods have the advantage of simplicity, which means they can be completed with relatively low costs and in relatively short time frames. However, judgmental forecasts are not typically based on data and do not typically involve any empirical analysis. As such, they may be prone to error and bias, depending on who is involved. That said, judgmental forecasts should rely on multiple participants, which may mean that the forecasts are somewhat more accurate. Of course, the accuracy of these forecasts depends greatly on how knowledgeable and insightful the participants are.

Salesforce Composite. In organizational markets, salespeople should know the needs and plans of their customers. They should have a good sense from their customers about how much their customers

expect to purchase in the coming period. As such, salespeople are often a good source of information for creating forecasts. Many companies, especially those with large salesforces, survey their salespeople and ask them whether they believe their customers will be buying more, less, or about the same amount in the coming periods. Salesforce composites should be interpreted guardedly, however. Despite their (hopefully) extensive customer knowledge, salespeople may not always provide objective forecasts. For one thing, good salespeople tend to be very optimistic. If so, some may overstate their expectations. For another, salespeople may suspect that their individual sales quotas or targets are based on sales forecasts. If so, they may have an incentive to downplay management expectations. Management should try to develop a feel for how their sales staff approach sales forecasting and make adjustments as necessary.

Jury of Executive Opinion. This judgmental forecasting method asks executives to provide their best guess of what sales will be in the coming period. Their opinions are compiled in some way and a forecast produced. How the opinions are compiled is up to the manager preparing the forecast; no rules exist for combining the executives' predictions. For example, some executives, based on position or past predictions, may be given more weight than others. Importantly, executives need not be strictly internal to the company preparing the forecast. Companies may poll customers, outside experts, industry leaders, and others in addition to their own corporate executives. A relatively broad range of inputs is generally desirable assuming that everyone asked can reasonably be expected to have some expert knowledge of the company and its prospects for the coming periods. However, if executive opinions of expected sales are very different from one another, then the manager preparing the forecast should be somewhat concerned about the accuracy of the forecast and should seek more information from the executives about why they made the predictions that they did.

Delphi Method. The Delphi method of forecasting is a somewhat more structured variation of the jury of executive opinion. The name derives from the city of Delphi in Greece. In ancient times it was the site of the temple of Apollo. People would travel to the temple to ask the priestess of Apollo, named Pythia, for her predictions of the future. She became known as the "Oracle of Delphi." Thus, the Delphi method refers to predictions of the future, which is what sales forecasting is all about.

The Delphi method begins with a panel or jury of knowledgeable executives, however, it utilizes an iterative approach to trying to reach a consensus among them. The executives, who are anonymous to one another, are asked to forecast sales and submit their opinions to the forecaster along with the reasoning behind their predictions. The forecaster collects opinions from the executives, and then summarizes them, usually with an average of the initial forecasts and some of the executives' reasons for their opinions. The summary is distributed back to the executives, who may revise their estimates and submit the revisions to the forecaster. The process continues until no more movement toward a consensus can be obtained. The Delphi method is obviously more involved than other judgmental methods and its iterative approach means it can take more time. However, its emphasis on consensus building along with well informed executives, often produce fairly accurate forecasts.

Simple Quantitative Forecasting Methods

Some even relatively simple forecasting methods rely on empirical data analyses to arrive at sales forecasts. We will discuss two of these simpler quantitative methods in this section.

Time Series Analysis. Let's begin by defining the term "time series." A time series is simply some variable reported at regular time intervals. In forecasting, the variable reported over time is usually either unit or dollar sales. Time series may be analyzed using highly complex statistical models or they may be analyzed using simple visual inspection. If the patterns that exist in past data hold true for the future data, then the forecaster should be able to project forward to predict future sales.

Look at Exhibit 1 to the right. It shows a hypothetical time series graph. Note that time is plotted on the horizontal axis and sales is on the vertical axis. In time series plots, this arrangement is always the case. As you look at the graph, do you see any patterns in the plot? If you look carefully, you should actually see more than one pattern occurring at the same time. That's because many time series, particularly of sales, tend to follow somewhat predictable recurring patterns. For example, many retailers have increasing sales in the months just preceding Christmas, followed by a decline in sales after the holiday season ends.

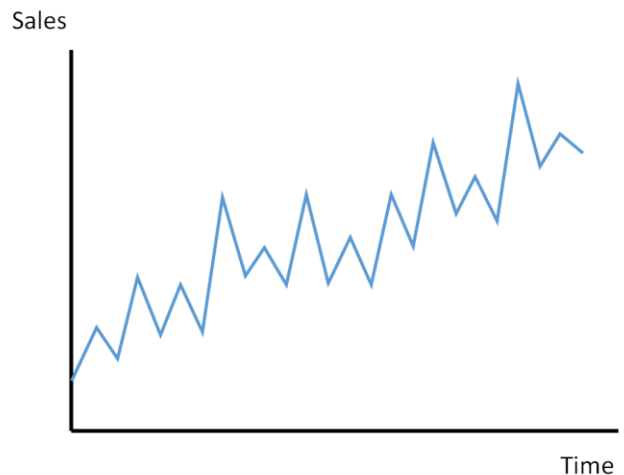


Exhibit 1. Time Series Plot

If retailer sales were graphed, this relatively predictable pattern would likely be easy to spot visually, assuming that monthly or even quarterly sales data were available. To better understand how to discern the patterns that may exist in time series data, it is useful to know the possible sources of variation in much time series data.

Look at Exhibit 2 to the right. It shows the same hypothetical time series graph as above. However, in this case it also shows the overall upward trend of sales over the time period covered by the time series. Thus, the first source of variation in time series data is called the *trend*. The trend shows the long term movement of the data. Trend is generally the longest term source of variation in time series data. For short term sales forecasts, trend is not especially useful. However, many sales forecasts attempt project sales out several periods. For such longer term forecasts, trend must be considered.

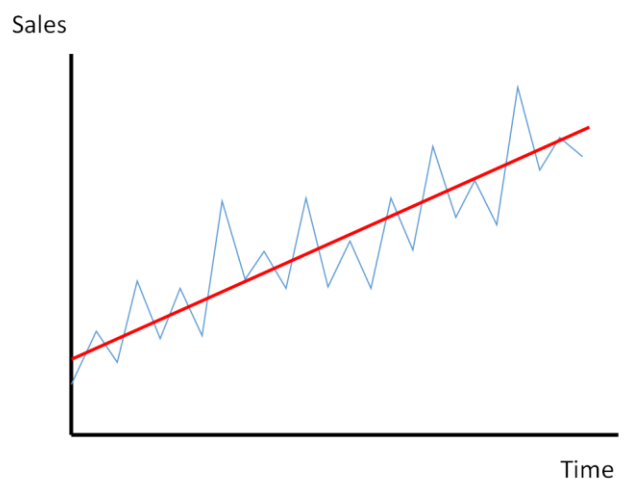


Exhibit 2. Time Series Plot showing Trend

Marketing managers, especially those in the United States, tend to be very focused on short term performance. Therefore, shorter term sources of time series variation are more useful for forecasting activities.

The Exhibit 3 to the right shows again the same time series data as on the previous page. The red line shows the second source of time series variation, called *cyclical*ity. When studying sales data, variation from cyclicality generally occurs with broad macroeconomic conditions. For example, the business cycle refers to the repeating ups and downs of gross domestic product over time. During periods of expansion, sales cyclicality may reflect this general pattern of increasing growth while during periods of recession, the general pattern may be one of slower growth or even contraction.

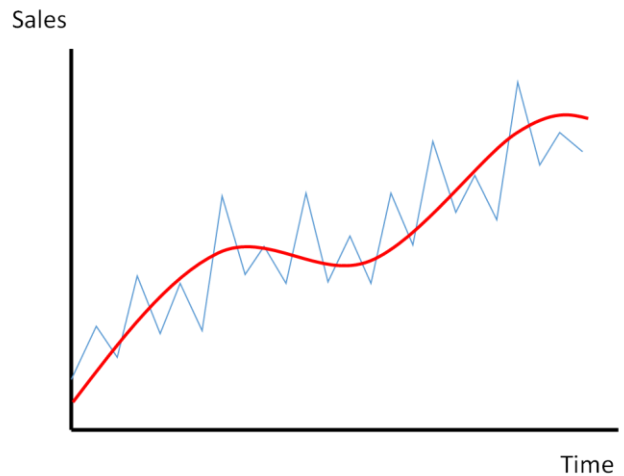


Exhibit 3. Time Series Plot showing Cyclical

A third source of variation in time series was described briefly earlier. This variation generally occurs over the course of a calendar year and generally captures the patterns that occur by month or by quarter. Look at Exhibit 4 to the right, which shows the same time series plot as the previous graphs. The red line shows the third source of variation, referred to as *seasonality*. The term reflects the fact that much intra-year variation in sales occurs by season. For example, as noted earlier, retailers often experience spikes in sales during the Christmas holidays and declines after the first start of the year. Seasonality is the shortest term source of variation in time series. Close visual, and as we will see in a moment, mathematical, inspection of seasonality is useful for making the types short term forecasts of interest to many marketing managers.

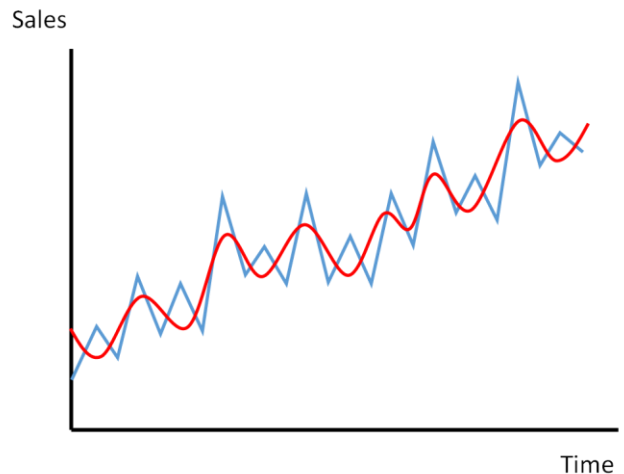


Exhibit 4. Time Series Plot showing Seasonality

The fourth and final source of variation in time series is simply random variance that occurs in every period. This is often referred to simply as “noise.” Because it is random, it cannot be predicted, meaning that even the best forecasts by any method cannot be perfect and will always have error..

Many mathematical approaches exist to remove some of all of the first three sources of variation from the time series in order to see the underlying structures of the patterns in sales. Some of these approaches are quite complex, while others are very simple but still effective in forecasting future period sales. In the following section, we look at one of the more common mathematical forecasting methods, the moving average.

Moving Averages. The concept of a moving average is really very simple, though they are best explained by example rather than by definition. Exhibit 5 shows a sample calculation of a moving average for hypothetical sales data. The calculations are very straightforward. Several points should be made about the example. First, notice that the example shows a three period moving average. Thus, the first entry in the moving average column begins with the third quarter of 2010. The first two spaces are blank because the moving average starts with the average of the first three quarters. Second, moving averages need not be limited to three periods. Mathematically, one can calculate moving averages to include as many time periods for data are available. Generally, however, moving averages rarely exceed five or seven periods. Third, time periods need not be quarters; they can be any time interval useful to the analysis. Quarterly and monthly data are common in moving averages of sales. Finally, moving averages (or time series generally for that matter) can be of any variable useful for analysis. Our emphasis here is on sales because it is the most commonly forecasted variable. However, variables such as event attendance or store traffic are also amenable to these analytical methods.

The reason moving averages are useful for forecasting is illustrated in Exhibit 6 on the following page. This exhibit shows the time series plot (blue line) of the data in Exhibit 5 as well as a plot of the moving average (red line). The main feature of the plots in Exhibit 6 is that the line of the moving average has less variance in it than the plot of the original sales data. In other words, the moving average plot is smoother than the plot of the original data. This illustrates what moving averages do; they smooth time series data, removing some of the variance and making the general direction of the time series easier to project forward.

Time Period	Sales	Three Period Moving Average	
2010 Q1	32,500	...	
2010 Q2	28,270	...	
2010 Q3	31,050	30,607	Average of 2010 Q1, Q2, and Q3
2010 Q4	30,990	30,103	Average of 2010 Q2, Q3, and Q4
2011 Q1	34,540	32,193	Average of 2010 Q3, Q4 and 2011 Q1
2011 Q2	26,780	30,770	
2011 Q3	32,460	31,260	and so on. . .
2011 Q4	31,770	30,337	
2012 Q1	34,760	32,997	
2012 Q2	30,920	32,483	

Exhibit 5. Moving Average Calculation

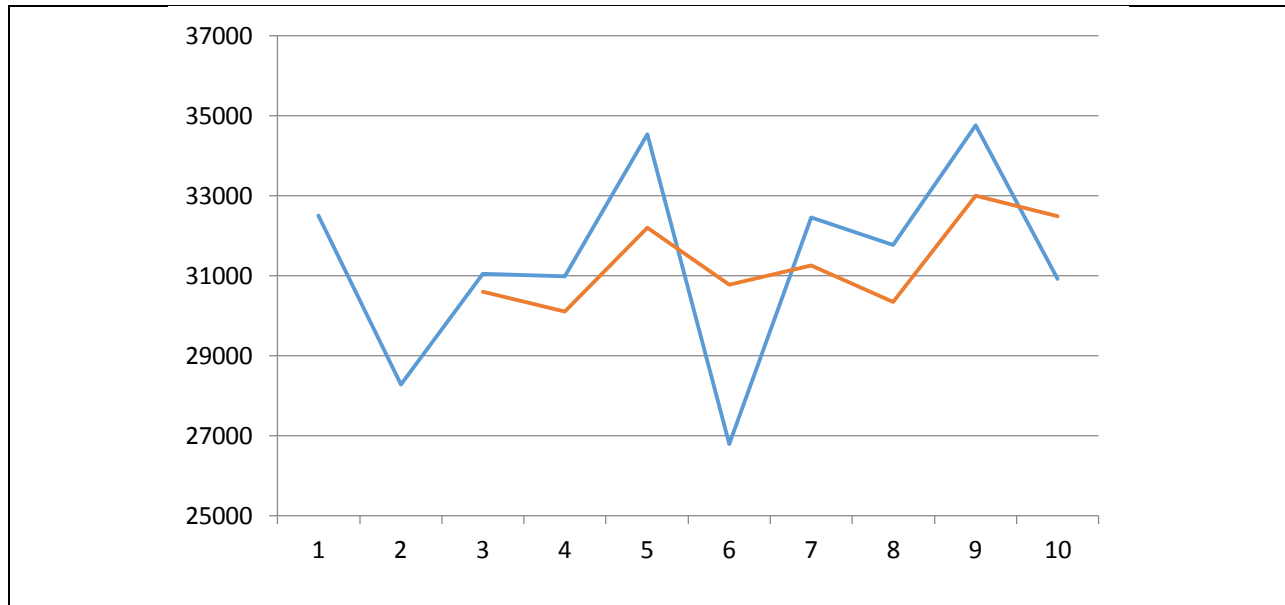


Exhibit 6. Plot of Time Series and Moving Average

The smoothing effect of moving averages can be seen with greater clarity in Exhibit 7 below. This plot shows real sales data from 1997 to 2011. The blue line shows the original time series and the orange line shows a five period moving average of the data. Notice how the five period moving average removes much of the variation from the original time series, especially seasonality. What remains is the trend, which is upward, and the cyclicality, which follows the broad wave shape of the curve.

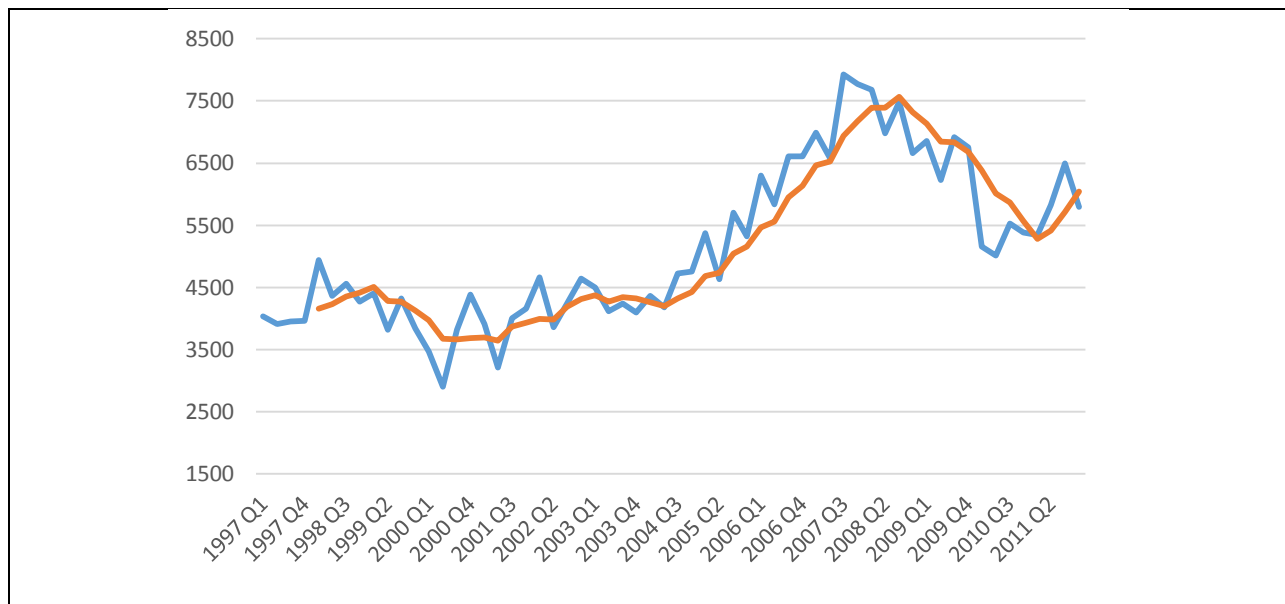


Exhibit 7. Time Series of Sales Data and Five Period Moving Average

To actually forecast using the moving average plot, the forecaster could simply extend the smoothed line out one quarter or more. The forecaster could visually adjust the angle based on where the moving average is relative to the trend and cyclical of the original time series. What is especially helpful regarding the data in Exhibit 7 is that the moving average appears to have moved into a period of upward average sales rather than being near the top or bottom of the curve. Predicting points of upturn or downturn is more difficult when using these simple forecasting techniques.

Importantly, forecasting using time series and moving averages can produce models that can predict with reasonable accuracy when upturns or downturns may occur, though their exact timing and magnitude can be difficult to pinpoint exactly. These models, called ARMA (autoregressive moving averages) and ARIMA (autoregressive integrated moving averages) models create functions that can closely reproduce even complex time series patterns, allowing forecasters to mathematically project future sales rather than relying on visual analysis only.