**What Is Demand Forecasting in Retail?**

**A Guide for Growing Businesses.**

Demand forecasting is a key component to every growing retail business. Without proper demand forecasting processes in place, it can be nearly impossible to have the right amount of stock on hand at any given time.

Too much merchandise in the warehouse means more capital tied up in inventory, and not enough could lead to out-of-stocks — and push customers to seek solutions from your competitors.

So, what is demand forecasting? And how is demand forecasting done in retail? Below, we’ll explain demand forecasting and how you can use it to support your retail business’ sustainable growth.

**Table of contents**

What is demand forecasting?

Why demand forecasting is important

Uses of demand forecasting

How is demand forecasting done?

Demand forecasting tips

How to calculate demand forecasting accuracy

**What is demand forecasting?**

Demand forecasting in retail is the act of [using data and insights](https://www.stitchlabs.com/blog/data-driven-retail-5-inventory-management-kpis-for-scalable-growth/) to predict how much of a specific product or service customers will want to purchase during a defined time period. This method of [predictive analytics](https://www.stitchlabs.com/press/stitch-labs-continues-automating-retail-launches-predictive-analytics/) helps retailers understand how much stock to have on hand at a given time.

### **What is demand forecasting in economics?**

Demand forecasting in economics is a bit different than how a retailer might use demand forecasting in business. So what do we mean by demand forecasting in economics, and how does that differ from retail?

In economics, analysts look at demand in the market as a whole, often for a particular industry or product category. In retail, you’ll look at the demand for YOUR products specifically. Demand forecasting in economics can (and should) inform forecasting in retail.

**What is demand forecasting in marketing?**

Demand forecasting in marketing is another component for retailers to consider. [Get your marketing and operations teams on the same page](https://www.stitchlabs.com/blog/marketing-and-operations-retail-success/) so that they can share calendars, priorities and initiatives and be proactive in planning. Retail ops can’t provide [inventory analytics](https://www.stitchlabs.com/retail-analytics-reporting/) for extra demand from a marketing campaign if they don’t know about it in the first place.

**Why demand forecasting is important?**

When explaining why demand forecasting is important, the answer spans across several areas of a retail business. One [Retail Systems Research report](https://www.sas.com/en_us/whitepapers/rsr-merchandising-omnichannel-impact-109055.html) found that nearly three-quarters of “winning” retailers rate demand forecasting technologies as “very important” to their business and their success.

**How does demand forecasting contribute to growing businesses?**

It mostly comes down to two things: becoming more cost-efficient and improving the customer experience.

How demand forecasting makes your business more cost-efficient

Almost every retail business is always looking for ways to cut costs. It’s one of the easiest ways to maximize your profits. When you implement a proper demand forecasting process to your business, you’re cutting costs in a few ways.

Firstly, you’re reducing the amount of capital you have tied up in unneeded inventory. And the less stock on hand you have, the lower your holding costs.

Secondly, you’re making sure you capitalize on every sale opportunity by not disappointing customers with out-of-stocks.

Those are the two most straightforward ways, but you can also use demand forecasting to operate a lean and agile business, only investing money in more stock when you need to. When you’ve forecasted demand, you can easily check in before the period’s over to see if you’re on target to hit your predicted sales. If you’re looking shy of your goal, you can [amp up marketing](https://www.stitchlabs.com/blog/growth-marketing-tactics-and-tips-for-retail-success/) and advertising. If it looks like you’ve underestimated, you could reorder or prep yourself to cross-promote a related product.

How demand forecasting enhances the customer experience

Another quick way to improve profits? Improve the customer experience. Rather than raising prices, focusing on the end user of the product can lead to customer loyalty and referrals.

Let’s go back to the most obvious: avoiding out-of-stocks that disappoint customers and lead them to your competitors. This is one of the most impactful ways to please customers.

Beyond simply having enough product to meet demand, you can also use forecasting to inform staffing decisions. While this is relevant to businesses needing [e commerce management](https://www.stitchlabs.com/integration-category/ecommerce-platforms/), it especially pertains to brick-and-mortar retailers. Customers who come to your store want to speak to an associate. And if no one’s there to help them, this can make a poor impression on shoppers. Even online sellers need to prep staff accordingly, especially during busy selling periods, so as not to delay [shipping and fulfillment](https://www.stitchlabs.com/resource/guides/simplifying-shipping-fulfillment/).

**Uses of demand forecasting**

As mentioned earlier, demand forecasting impacts many areas of your retail business. Here are just a few use cases of demand forecasting for rapidly growing businesses needing [multichannel management](https://www.stitchlabs.com/multichannel-inventory-management/):

* Prepare accurate budgets and financial planning
* Make informed purchasing decisions
* Implement [purchase order](https://www.stitchlabs.com/purchase-orders/) automations to avoid stock issues
* Gain a thorough, comprehensive understanding of your business
* Anticipate staffing needs
* Grow sustainably
* Measure progress towards business and sales objectives
* Streamline production process
* [Plan advertising and marketing campaigns](https://www.stitchlabs.com/resource/guides/developing-scalable-marketing-strategies/) and budgets
* [Enhance the customer experience](https://www.stitchlabs.com/resource/guides/creating-seamless-customer-experience/) (avoid out-of-stocks, backorders, late shipments, etc.)
* Resourcing and project management

**How is demand forecasting done, accurately?**

Rather than asking “how is demand forecasting done?”, retailers should ask “how is demand forecasting done*most accurately?*” There are many flaws to every approach to estimating demand and forecasting. Even though we can’t predict the future perfectly, using established methods can help you be more successful in your forecasting practices.

Demand forecasting is done most accurately when a business considers both internal and external [data](https://www.stitchlabs.com/blog/five-data-points-that-youre-missing-with-spreadsheets/). Internal metrics may include historical sales numbers, ad spend, and website or foot traffic. Externally speaking, you’re looking at factors like industry or consumer trends, the weather, and even your competitors.

To best explain demand forecasting, it’s helpful to look at the different methods. Some of the most common demand forecasting techniques include:

* Qualitative forecasting
* Time series analysis
* Causal model

**Qualitative forecasting**

This type of forecasting is when a business anticipates demand based on qualitative data. Qualitative data sources could include industry experts and/or consultants, employees, focus groups, and competitive analysis, to name a few. Often, this data is subjective and based on intuition rather than hard numbers or facts.

* Market research
* Delphi Method
* Expert opinion
* Focus groups
* Historical analogy
* Panel consensus
* Surveys

Recommended for: businesses that have limited historical data; new product launches (especially if there’s no other product like it on the market); instances where the previous period is believed to differ drastically from the planned period (for example, the [Tickle Me Elmo frenzy](https://www.timesunion.com/life/article/Remember-the-Tickle-Me-Elmo-craze-Hot-10781630.php) during the 1996 holiday season)

**Time series analysis**

The time series analysis is a more quantitative approach to demand and forecasting. Rather than expert opinions and “soft” data inputs, a time series analysis uses exact numbers as the basis for forecasting demand. It’s a more mathematical approach to forecasting which uses numerical inputs and trends.

Other quantitative forecasting methods include:

* The indicator approach
* Econometric modeling
* Trend analysis
* Seasonal adjustment
* Decomposition
* Graphical methods
* Life cycle modeling

Recommended for: retailers that have plenty of past sales data (especially if this data reveals year-over-year trends); seasonal items; [seasonal selling](https://www.stitchlabs.com/blog/turn-seasonal-shoppers-lifelong-brand-advocates/) periods; identifying cyclical sales trends

**Causal model**

The causal model accounts for demand forecasting factors that may change predicted demand. Demand forecasting factors are both controllable and uncontrollable:

|  |  |
| --- | --- |
| Controllable demand factors | Uncontrollable demand factors |
| Marketing, sales and promotions | Weather |
| Price | Politics |
| Visual merchandising | Trends |
| Location | Competitors |
|  | Economic and socioeconomic conditions |
|  | Seasonality |

Because the causal method of forecasting accounts for so many variables, it’s also a more complex approach. Some of the factors, like the weather, can’t be predicted as accurately as you might like. This includes a part guesswork, part data-driven approach to forecasting — and a lot of trust in your intuition.

Recommended for: data-driven retailers with lots of metrics; forecasting by specific product, category or SKU; retailers in volatile markets; multi-channel businesses with a diverse customer base; forecasting in association with marketing/advertising campaigns and promotions

Demand forecasting tips

Demand forecasting is half art, half science. The best approach is to account for qualitative and quantitative data, internal and external variables, and controllable and uncontrollable factors. Many assumptions must be made, as well as “guesstimations” based off your experiences.

That being said, there are a few tips for demand forecasting that you can apply to ensure you’re doing it properly:

Establish a baseline: This should be the first task on your list, aside from establishing a goal or hypothesis that you’ll want to achieve or answer with your forecast. Without having a baseline of data, you’re solely going off of third-party information.

Preserve your data: Because using your own data is so valuable in demand forecasting, you’ll also need to ensure the data is clean and accurate. [Centralize your inventory information](https://www.stitchlabs.com/why-centralized-inventory-management-is-the-key-to-profitability-in-an-omnichannel-world/) so that everything is synced and in a single location, and you’ll mitigate discrepancies.

Invest in the right tools: Without the right tools, demand forecasting can be a tedious, manual process. Find the right [inventory management software](https://www.stitchlabs.com/inventory-management-software/) that integrates with your accounting, point-of-sale and other tools for the most comprehensive look at your business.  
  
It’s not always clear [what to look for in an inventory system](https://www.stitchlabs.com/resource/ebooks/how-to-choose-inventory-management-software-for-retail-success/), so we created a guide to help.

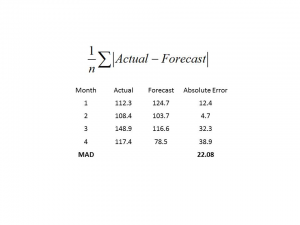
How to calculate demand forecasting accuracy

It’d be remiss to explain demand forecasting without also describing how to calculate demand forecasting accuracy. After all, demand forecasting can be done by almost anyone — but it’s not always done accurately. And if your forecast is inaccurate, then you risk making majorly impactful business decisions based off the wrong information.

To calculate demand forecasting accuracy, many retailers look at the Mean Absolute Deviation (MAD) and Mean Absolute Percent Error (MAPE).

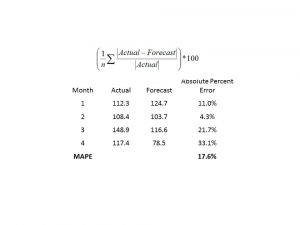
**MAD =**

MAD is the average difference between the actual demand and forecasted demand. To calculate MAD, you’ll subtract the forecasted demand from the actual demand. You can then average this number over several time periods to find out your overall MAD.



**MAPE =**

MAPE measures the rate of accuracy of your forecast and is calculated by subtracting the forecasted demand from the actual demand, and then dividing that number by the actual demand. To get the percentage, multiply by 100. Again, you’ll calculate this for multiple time periods and determine the average to find out your MAPE.



Business forecasting is essential for the survival for companies of all sizes. The building block used by forecasters is historical data or the past performance of the business to predict future results. Regression analysis is a statistical technique used to find relationships between independent and dependent variables. Regression analysis uses historical data and observation to predict future values.

**Historical Data**

Business forecasting by its very nature uses historical data to forecast future performance of the company. Historical data includes your company's financial statements, client invoices and any information you believe has relative predictive value to the future success of your company. Historical data doesn't have to solely come from your company; it can also be historical macroeconomic data, such as the Consumer Confidence Index, interest rates, housing starts or any other economic variable you believe has an affect on your business based on your observations and business experience.

**Regression Analysis**

Regression analysis applies to almost any field. In general, regression analysis identifies relationships based on independent and dependent variables. For example, a dependent variable is your company's sales and an independent variable may be interest rates. Governments and businesses use regression analysis as a predictive tool for forecasting purposes. Regression analysis measures the strength or correlation between the dependent and independent variables. For the uninitiated, regression analysis may become complex particularly with the addition of multiple independent and dependent variables, requiring sophisticated programs and analysis.

**Forecasting**

Business forecasting using historical data is tricky business. Past performance is not necessarily a good indicator of future performance. In other words, you shouldn't forecast more than 10 percent in sales growth in the next period because your company's sales grew more than 10 percent in three prior periods. This is where regression analysis comes in. By identifying the relationship between the dependent variable or your company's sales with independent variables, you may gain a better insight as to the factors that determine your company's sales growth. Generally speaking, the predictive power of regression analysis improves the longer the time period used to build your forecast.

**Insight**

Using historical data and regression analysis has its limitations in business forecasting. For example, a significant correlation between the independent and dependent variable does not necessarily indicate a cause and effect relationship. In some cases, the common linkage may be because of a sequence of events. If you are considering whether to incorporate regression analysis as a forecasting tool for your business, there are several user-friendly statistical programs available to help you perform regression analysis, which includes identifying statistical errors that may affect your forecast.

**TIME SERIES ANALYSIS**

AR, MA, ARMA, and ARIMA models are used to forecast the observation at (t+1) based on the historical data of previous time spots recorded for the same observation. However, it is necessary to make sure that the time series is stationary over the historical data of observation overtime period. If the time series is not stationary then we could apply the differencing factor on the records and see if the graph of the time series is a stationary overtime period.

**ACF (Auto Correlation Function)**

Auto Correlation function takes into consideration of all the past observations irrespective of its effect on the future or present time period. It calculates the correlation between the t and (t-k) time period. It includes all the lags or intervals between t and (t-k) time periods. Correlation is always calculated using the Pearson Correlation formula.

**PACF(Partial Correlation Function)**

The PACF determines the partial correlation between time period t and t-k. It doesn’t take into consideration all the time lags between t and t-k. For e.g. let's assume that today's stock price may be dependent on 3 days prior stock price but it might not take into consideration yesterday's stock price closure. Hence we consider only the time lags having a direct impact on future time period by neglecting the insignificant time lags in between the two-time slots t and t-k.

**How to differentiate when to use ACF and PACF?**

Let's take an example of sweets sale and income generated in a village over a year. Under the assumption that every 2 months there is a festival in the village, we take out the historical data of sweets sale and income generated for 12 months. If we plot the time as month then we can observe that when it comes to calculating the sweets sale we are interested in only alternate months as the sale of sweets increases every two months. But if we are to consider the income generated next month then we have to take into consideration all the 12 months of last year.

So in the above situation, we will use ACF to find out the income generated in the future but we will be using PACF to find out the sweets sold in the next month.

**AR (Auto-Regressive) Model**

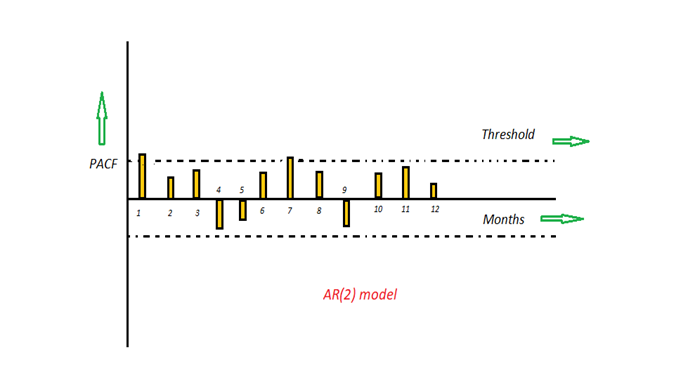


Image by Author

The time period at t is impacted by the observation at various slots t-1, t-2, t-3, ….., t-k. The impact of previous time spots is decided by the coefficient factor at that particular period of time. The price of a share of any particular company X may depend on all the previous share prices in the time series. This kind of model calculates the regression of past time series and calculates the present or future values in the series in know as Auto Regression (AR) model.

**Yt = β₁\* y-₁ + β₂\* yₜ-₂ + β₃ \* yₜ-₃ + ………… + βₖ \* yₜ-ₖ**

Consider an example of a milk distribution company that produces milk every month in the country. We want to calculate the amount of milk to be produced current month considering the milk generated in the last year. We begin by calculating the PACF values of all the 12 lags with respect to the current month. If the value of the PACF of any particular month is more than a significant value only those values will be considered for the model analysis.

For e.g in the above figure the values 1,2, 3 up to 12 displays the direct effect(PACF) of the milk production in the current month w.r.t the given the lag t. If we consider two significant values above the threshold then the model will be termed as AR(2).

**Python Code**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11 | # AR example  from statsmodels.tsa.ar\_model import AutoReg  from random import random  # contrived dataset  data = [x + random() for x in range(1, 100)]  # fit model  model = AutoReg(data, lags=1)  model\_fit = model.fit()  # make prediction  yhat = model\_fit.predict(len(data), len(data))  print(yhat) |

**MA (Moving Average) Model**

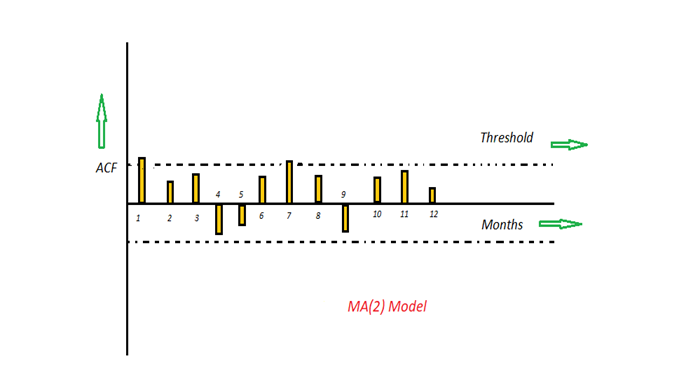


Image by Author

The time period at t is impacted by the unexpected external factors at various slots t-1, t-2, t-3, ….., t-k. These unexpected impacts are known as Errors or Residuals. The impact of previous time spots is decided by the coefficient factor α at that particular period of time. The price of a share of any particular company X may depend on some company merger that happened overnight or maybe the company resulted in shutdown due to bankruptcy. This kind of model calculates the residuals or errors of past time series and calculates the present or future values in the series in know as Moving Average (MA) model.

**Yt = α₁\* Ɛₜ-₁ + α₂ \* Ɛₜ-₂ + α₃ \* Ɛₜ-₃ + ………… + αₖ \* Ɛₜ-ₖ**

Consider an example of Cake distribution during my birthday. Let's assume that your mom asks you to bring pastries to the party. Every year you miss judging the no of invites to the party and end upbringing more or less no of cakes as per requirement. The difference in the actual and expected results in the error. So you want to avoid the error for this year hence we apply the moving average model on the time series and calculate the no of pastries needed this year based on past collective errors. Next, calculate the ACF values of all the lags in the time series. If the value of the ACF of any particular month is more than a significant value only those values will be considered for the model analysis.

For e.g in the above figure the values 1,2, 3 up to 12 displays the total error(ACF) of count in pastries current month w.r.t the given the lag t by considering all the in-between lags between time t and current month. If we consider two significant values above the threshold then the model will be termed as MA(2).

**Python Code**

We can use the ARIMA class to create an MA model and setting a zeroth-order AR model. We must specify the order of the MA model in the order argument.

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11 | # MA example  from statsmodels.tsa.arima.model import ARIMA  from random import random  # contrived dataset  data = [x + random() for x in range(1, 100)]  # fit model  model = ARIMA(data, order=(0, 0, 1))  model\_fit = model.fit()  # make prediction  yhat = model\_fit.predict(len(data), len(data))  print(yhat) |

**ARMA (Auto Regressive Moving Average) Model**

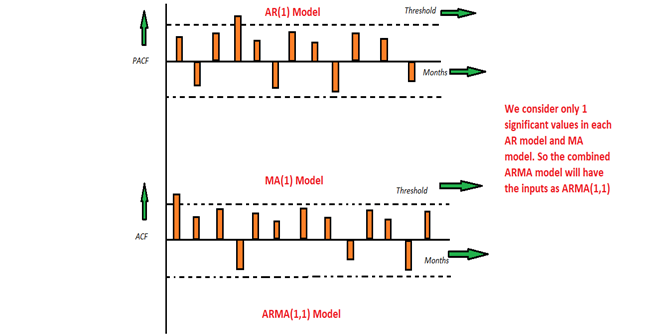


Image by Author

This is a model that is combined from the AR and MA models. In this model, the impact of previous lags along with the residuals is considered for forecasting the future values of the time series. Here β represents the coefficients of the AR model and α represents the coefficients of the MA model.

**Yt = β₁\* yₜ-₁ + α₁\* Ɛₜ-₁ + β₂\* yₜ-₂ + α₂ \* Ɛₜ-₂ + β₃ \* yₜ-₃ + α₃ \* Ɛₜ-₃ +………… + βₖ \* yₜ-ₖ + αₖ \* Ɛₜ-ₖ**

Consider the above graphs where the MA and AR values are plotted with their respective significant values. Let's assume that we consider only 1 significant value from the AR model and likewise 1 significant value from the MA model. So the ARMA model will be obtained from the combined values of the other two models will be of the order of ARMA(1,1).

**Python Code**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11 | # ARMA example  from statsmodels.tsa.arima.model import ARIMA  from random import random  # contrived dataset  data = [random() for x in range(1, 100)]  # fit model  model = ARIMA(data, order=(2, 0, 1))  model\_fit = model.fit()  # make prediction  yhat = model\_fit.predict(len(data), len(data))  print(yhat) |

**ARIMA (Auto-Regressive Integrated Moving Average) Model**

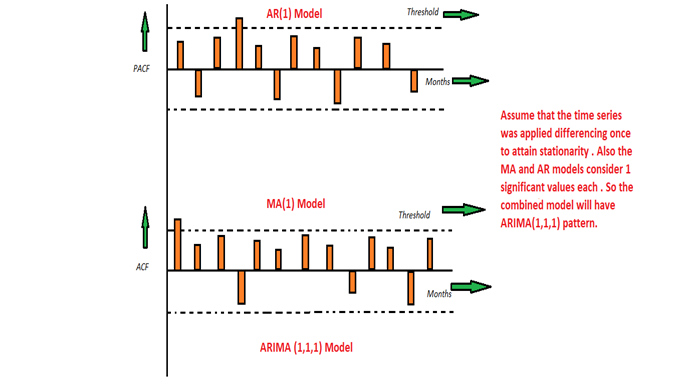


Image by Author

We know that in order to apply the various models we must in the beginning convert the series into Stationary Time Series. In order to achieve the same, we apply the differencing or Integrated method where we subtract the t-1 value from t values of time series. After applying the first differencing if we are still unable to get the Stationary time series then we again apply the second-order differencing.

The ARIMA model is quite similar to the ARMA model other than the fact that it includes one more factor known as Integrated( I ) i.e. differencing which stands for I in the ARIMA model. So in short ARIMA model is a combination of a number of differences already applied on the model in order to make it stationary, the number of previous lags along with residuals errors in order to forecast future values.

Consider the above graphs where the MA and AR values are plotted with their respective significant values. Let's assume that we consider only 1 significant value from the AR model and likewise 1 significant value from the MA model. Also, the graph was initially non-stationary and we had to perform differencing operation once in order to convert into a stationary set. Hence the ARIMA model which will be obtained from the combined values of the other two models along with the Integral operator can be displayed as ARIMA(1,1,1).

**Python Code**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11 | # ARIMA example  from statsmodels.tsa.arima.model import ARIMA  from random import random  # contrived dataset  data = [x + random() for x in range(1, 100)]  # fit model  model = ARIMA(data, order=(1, 1, 1))  model\_fit = model.fit()  # make prediction  yhat = model\_fit.predict(len(data), len(data), typ='levels')  print(yhat) |

Conclusion :

All these models give us an insight or at least close enough prediction about any particular time series. Also, it depends on the users that which model perfectly suffices their needs. If the chances of error rate are less in any one model compared to other models then it's preferred that we choose the one which gives us the closest estimation.