Regression for Predictive Analytics

Finding the trend for Business

Photo by Franki Chamaki on Unsplash

Predictive Analysis adapts variety of statistical techniques from data mining, predictive modelling, and machine learning that analyze current and historical data to make predictions about future or otherwise unknown events. This type of analysis uses a pool of stored information to generate trends and forecasting to understand “what happens next”.

Understanding future behaviors and data trends allows for a reliable and repeatable was to engage and plan for work, home, or school.

How to use Predictive Analysis

There are several methods for predictive analysis. All ways to generate predictions of future behaviors require a pool of past or current data. Using predictive modeling and data mining among other techniques, proper trends can be generated. The current top method for predictive analysis is regression. Any model that creates a trend line for estimation based on data stored in a network is a good choice for analysis. Linear regression is the most used predictive analysis method. Excel with a sample dataset used to show predictive analysis with linear regression.

The **benefit** of **regression** analysis is that it can be used to understand all kinds of patterns that occur in data. These new insights may often be very valuable in understanding what can make a difference in your business.

Applying Prediction in a Grocery Store

An easy example of a use case for predictive analytics is estimating future needs for stock in a grocery store. Stores have a supply chain to have goods or services to sell customers. There needs to be enough stock to provide for demand without excess to have an efficient business. An example is selling drinks in a store, where there is past data of units purchased by date and type. Using this data, a report showing the past is usable information to show history of sales. Creating a trendline based on regression using the historical data continuing into the future predicts the demand and creates a basis for selecting next steps to evaluate the supply for obtaining drinks to sell at a grocery store.

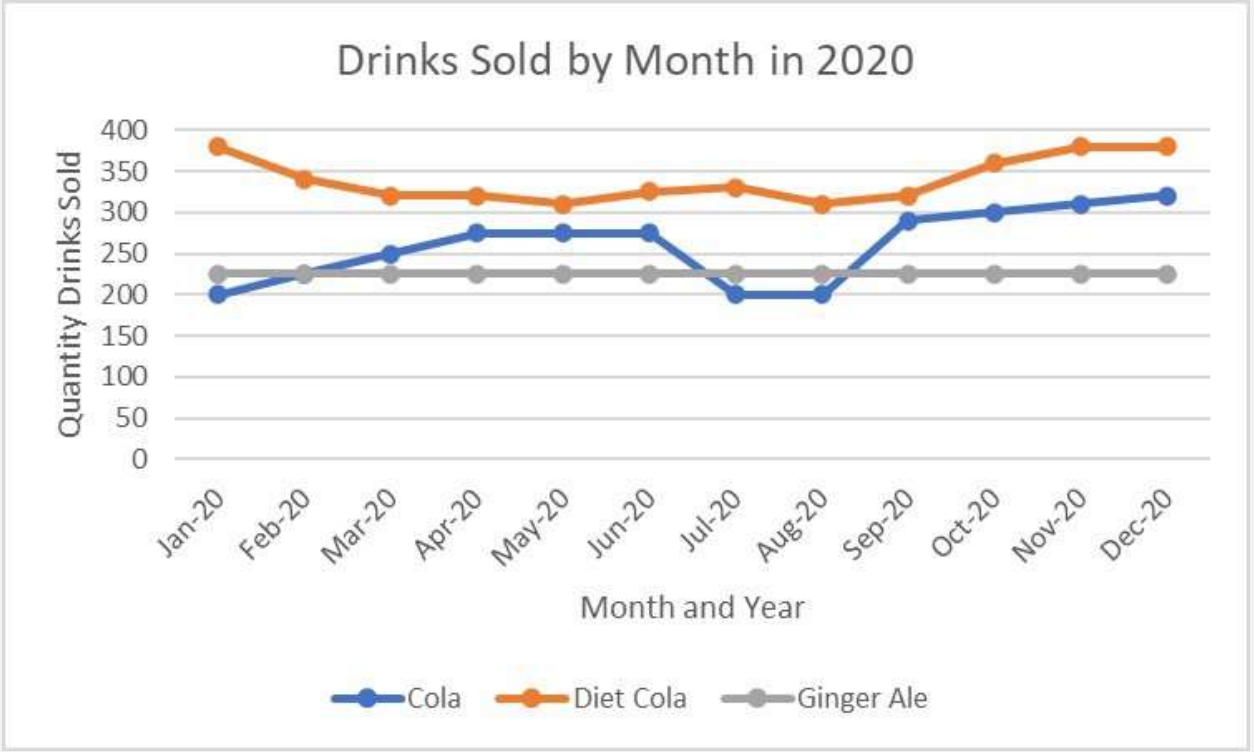

Figure 1

From Figure 1, we can see a visual representation of drinks sold by type during 2020 at a grocery store. Each dot represents a discrete amount sold in a month with a line connecting the dots. This is a representation of past data.

Prediction from Data using Linear Regression Model

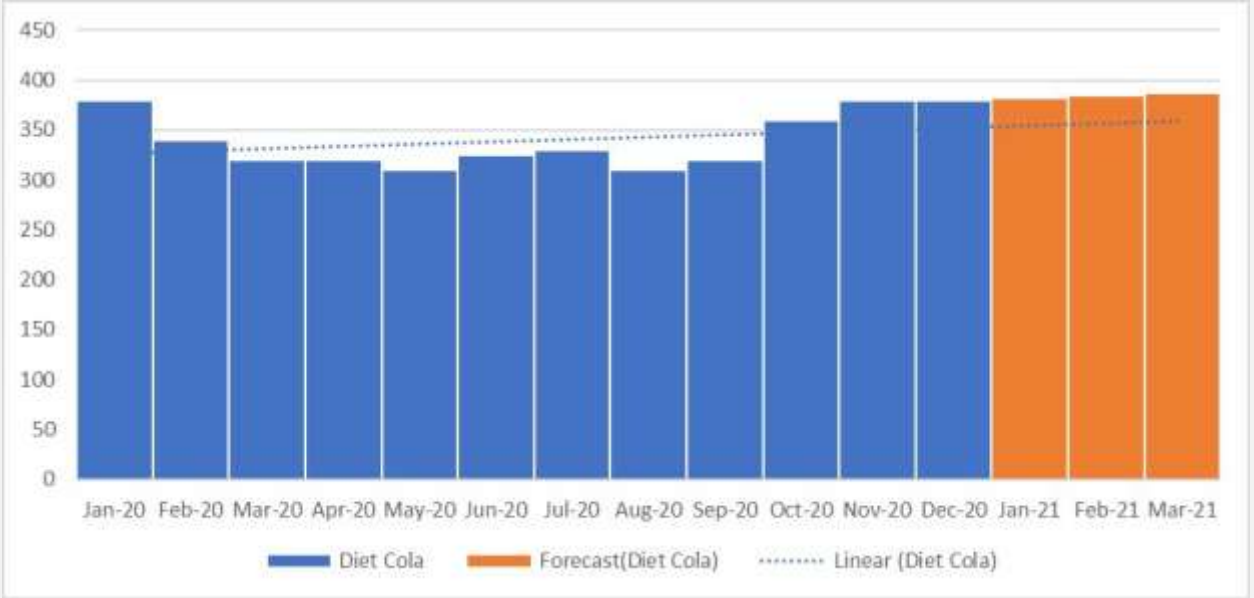

Figure 2

From Figure 2, we can see a visual representation of Diet Cola sold during 2020 at a grocery store with prediction of sales for the first three months of 2021. Each bar is one month, with blue as past and orange future values, the line is the regression model. This shows an increase in sales is expected, so more Diet Cola should be ordered in 2021 than 2020.

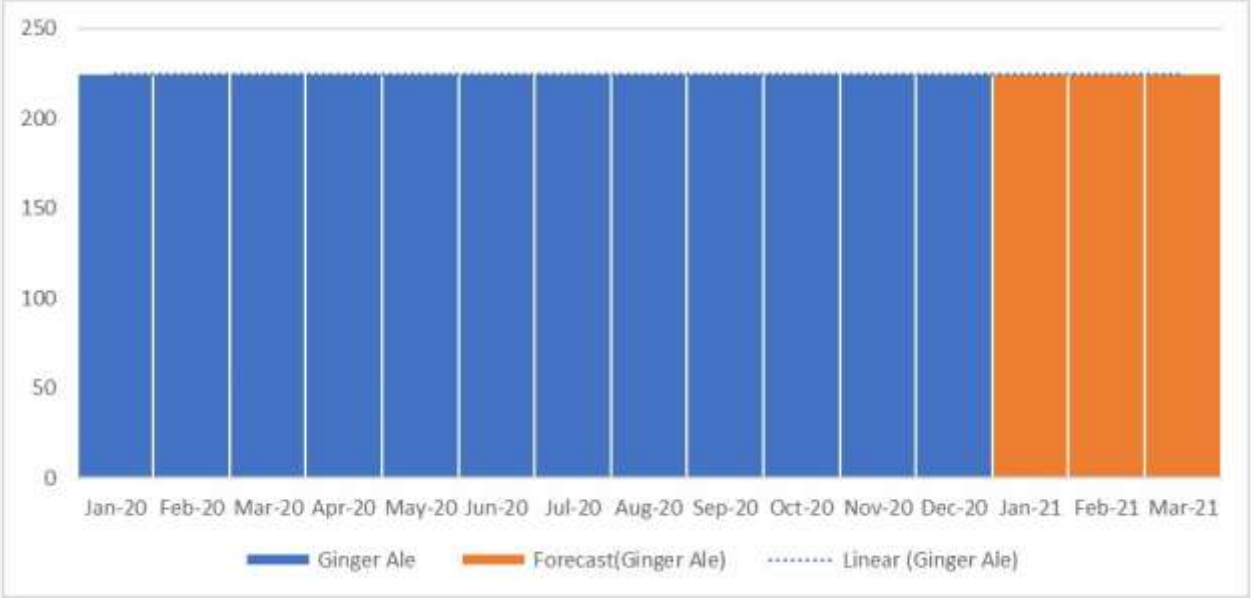

Figure 3

From Figure 3, we can see a visual representation of drinks sold that are Ginger Ale by month and year. Each bar is one month, with blue as past and orange future values, the line is the regression model as a flat line continuing the trend of the same amount of Ginger Ale sold each month based on the previous year.

Using visualizations, the historical data is clear and concise showing the amount of drinks sold by type and when. The predictive model, linear regression, shows the probable future amount of sales by creating a linear trend line and calculating three months based on data from the previous year. This allows the grocery store to expect demand for product and order to ensure it can supply drinks to the customers. This allows for a more efficient stock management by predicting quantity to plan for supply to meet predicted demand.

Originally published:

https://medium.com/codex/regression-for-predictive-analytics-9bf96deef3f5