# New Product Forecasting of Appliance and Consumables: Bass Model

**Keval Babu**

**Abstract**

**Drinkworks, a joint venture by Anheuser-Busch and Keurig Green Mountain Inc., has created an in-home alcohol drink system. The challenge was predicting demand throughout the product's life-cycle. As a new organization, Drinkworks needed a systematic demand planning tool for baseline strategic and operational forecasts, which would aid in sales and operations, production planning, and material resource planning. This thesis focuses on selecting mathematical models to forecast demand, particularly using the Bass model for the initial launch phase despite limited market knowledge. It also details the methodology to forecast pod consumption, including the average consumption rate per appliance, cumulative appliances sold to retailers, cluster analysis, and appliance forecast.**

**Keywords: Bass model, pods, SARIMA, SMA, associative model, forecasting, p, q,m, DP, SKU**

## I. INTRODUCTION

Based in Bedford, Massachusetts, Drinkworks specializes in creating an in-home alcohol drink system capable of crafting various alcoholic beverages, including beers, cocktails, and mixers. This venture is a collaboration between Anheuser-Busch InBev, one of the globe's largest brewers, and Keurig Green Mountain, Inc., a prominent beverage systems company. Their primary goal is to dominate the in-home alcoholic beverage systems market by leveraging the combined expertise of both partners. Figure 1.1 depicts the Drinkworks logo.

**Figure 1. Drinkworks' logo**

## II. PRODUCT OVERVIEW

The Drinkworks in-home alcohol drink system comprises two main sub-categories: the appliance and consumables. The appliance is an automated machine designed to brew a variety of alcoholic beverages using chilled water, Carbon dioxide ($CO_2$) gas, and disposable pods. Each single-serving pod contains a drink concentrate that determines the type of drink produced and consists of alcohol and flavoring agents. The $CO_2$ usage varies depending on the beverage type.

Initially, the appliance will be marketed as a single SKU (Stock Keeping Unit), while the pods will be available in different SKUs, each representing a unique type of alcoholic drink. Drinkworks' business model mirrors Gillette's Razor-Razorblade strategy, where the independent appliance is sold at cost, and the consumable pods generate profit. The process flow for creating a drink using the Drinkworks appliance and pods is illustrated in Figure2.

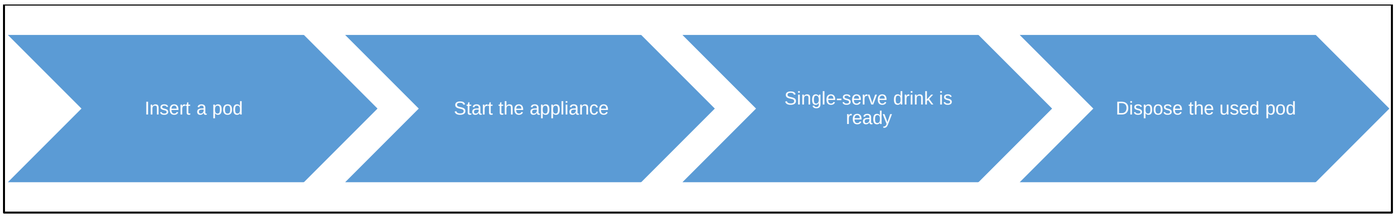


**FIGURE 2: DRINK MAKING PROCESS FLOW**

## III. APPROACH

The primary aim of this industry-focused project was to develop a Demand Planning (DP) software tool to establish a baseline demand forecast for Drinkworks' new "in-home alcohol drink system." The DP tool should generate both strategic and operational demand forecasts, serving as a foundation for sales and operations planning, production planning, and resource planning. The project scope was confined to forecasting demand for the appliance and consumable pods, which are essential components of the product.

The functional requirements for the DP software tool included:

1. Selection and evaluation of an appropriate mathematical model to forecast demand.
2. Capability to exchange data via CSV (comma-separated values) or MS Excel files and integration with an online database.
3. A graphical user interface to review and interact with forecast results.

Drinkworks intends to introduce their new product through two sequential phases: (i) the Pilot phase – a 5-month period designed to assess the feasibility of the product and gather initial sales data, and (ii) the US national launch – a nationwide rollout across selected cities. Figure 3 provides a summary of the project timeline.

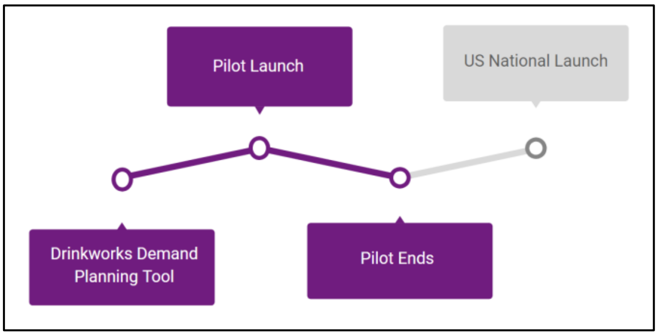


**FIGURE3: PROJECT LAUNCH PLAN**

## IV. METHODOLOGY

Drinkworks' in-home alcohol drink system represents a novel product in an uncharted market, lacking historical data. The forecasting approach was segmented into two product sub-categories:
(1) Appliance forecast – to predict the demand for standalone devices.
(2) Pods forecast – to anticipate the consumption of the dependent consumables.

(1) Appliance Forecast

In selecting the appropriate forecasting model to predict product demand, several factors were taken into account [3], including the context of the forecast, available historical data, benefit-to-cost ratio, degree of forecasting accuracy, forecasting time period, and the product's life-cycle stage. Based on these criteria, three distinct forecasting models were chosen:

(a) Bass Model – to provide a strategic forecast for the pilot and US national launch.

(b) Simple Moving Average Model (MA) – to generate operational forecasts during the pilot and US national launch.

(c) Seasonal Autoregressive Integrated Moving Average Model (SARIMA) – to produce more accurate operational forecasts during the US national launch.

Figure 4 illustrates the application of various forecasting models according to the product launch phase.

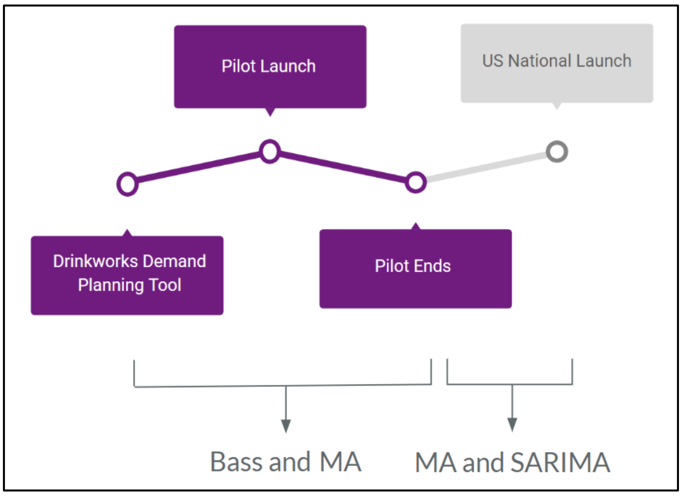


**Figure 4. Different forecasting models based on product launch phase**

(a) The Bass model

A widely utilized and significant diffusion model is employed to establish the initial life-cycle sales curve of a product. This model is specifically designed to predict the adoption and diffusion of a new product in markets where historical data is unavailable. Utilizing three parameters, the Bass model forecasts sales over a set period [4]. Consequently, it was chosen to create an initial strategic forecast for the appliances, given their novelty and lack of historical data.

(b) Simple Moving Average Model

The simple moving average (SMA) model is a straightforward time series model used to predict future demand based on historical data. Mathematically, the SMA model forecasts demand for the next period by calculating the arithmetic average of N recent observations, where N represents the number of historical data points considered. The SMA model effectively balances accuracy and complexity, providing stable forecasts [5]. However, it does not account for seasonality, trends, cycles, or irregular patterns.

Given the limited sales data during the early stages of the product launch, the SMA model will be employed to generate operational forecasts for the appliance. This model is expected to provide reasonable short-term forecasts for subsequent periods [1].

(c) Seasonal Autoregressive Integrated Moving Average

Unlike the simple moving average, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model considers the four key components of demand: trend, seasonality, cycle, and irregular patterns when forecasting future demand. This time series model utilizes historical sales data to make predictions. With adequate historical sales data, SARIMA can effectively incorporate these demand components into the product forecast. SARIMA identifies the best fitting model and its corresponding parameters through various estimation methods [5]. It performs optimally when the new product has achieved a steady state [3].

The SARIMA model will be employed to generate operational forecasts for the appliance after gathering one and a half years of sales data. It will be utilized to predict demand for the appliance once they have entered their maturity phase [2].

(2) Pods Forecast

The use of an associative model (causal model) is the optimal approach for forecasting the demand for dependent commodities based on the forecast of independent goods. An associative model predicts demand based on established correlations. Essentially, a linear regression analysis is used to establish a correlation between dependent and independent variables. This associative relationship can then be utilized to forecast the dependent variable [6].

In this project, the pods are dependent consumables whose consumption relies on the number of appliances sold. An associative model will be developed to forecast pod demand based on four main parameters: (i) the cumulative number of appliances sold to retailers, (ii) the appliance forecast, (iii) the percentage of active appliances, and (iv) the average pod consumption per appliance. The relationship established among these parameters will be used to accurately forecast pod consumption.

## V. Literature Review

Marketing researchers have frequently employed diffusion models to predict the demand for new products based on the estimated product life-cycle curve. These models have played a crucial role in guiding strategic decisions for new product launches [7].

In 1962, E.M. Rogers introduced the theory of diffusion of innovation, which defines two key concepts:
(a) Diffusion – the extent to which a new product spreads through a market from its creation to end-user adoption.
(b) Adoption – the process a potential user undergoes from learning about the new product to eventually purchasing it.

The theory categorizes new product adopters into two main groups based on their timing and motivation for adoption:
(i) Innovators – individuals who are the first to adopt a new product due to their curiosity and venturesome nature.
(ii) Imitators – individuals who delay adoption and base their decision on the experiences of prior adopters. Imitators are further divided into four sub-groups: early adopters, early majority, late majority, and laggards [4] [8].

Figure 5 illustrates the product diffusion curve, highlighting when different categories of adopters make their purchases.

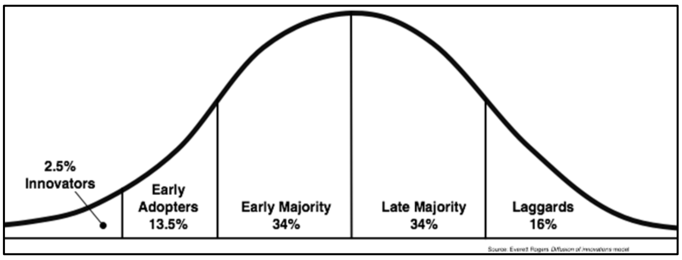


**Figure 5: Product diffusion curve [5]**

Among the most prevalent and recognized diffusion models in marketing is the Bass model [9], named after its creator, the esteemed marketing science professor Frank M. Bass. Over the years, the Bass model has garnered a reputation for accurately delineating and approximating the diffusion curve for new products and technologies [10].

The four key features of the Bass model are as follows [4] [11]:
(a) Adequate for predicting the initial purchase of a new product with no existing competing product in the market.
(b) Provides a reliable baseline long-term forecast for a new product.
(c) Requires three basic parameters to estimate the product life-cycle sales curve, which can be obtained through two methods:
• Analogous product – utilizing historical sales data or industry data from a similar new product launched previously.
• Early sales data – deriving initial sales data post-launch of the new product.
(d) Answers the question, “How and when will customers adopt the new product?”

The Bass model operates under several key assumptions [4] [11]:
(a) The diffusion process is twofold: customers either adopt or wait to adopt the new product.
(b) The model assumes a maximum number of potential adopters, ultimately predicting that all adopters will purchase the new product.
(c) It does not account for repeat or replacement purchases.
(d) The innovation factor is independent of the effect of substitutes.
(e) The adoption rate is independent of the product’s price.

The Bass model uses three input parameters to forecast a product's life-cycle sales curve. These parameters are:

(a) p - Coefficient of innovation (probability of adoption by innovators)
(b) q - Coefficient of imitation (probability of adoption by imitators)
(c) m - Market potential of the new product (total number of potential adopters)

The model determines the number of adopters at any given time based on the innovation and imitation effects. The innovation effect is calculated by multiplying the coefficient of innovation by the remaining market potential. Similarly, the imitation effect is the product of the coefficient of imitation, cumulative adopters up to the previous time period, and remaining market potential.

*The general equation of the Bass model is explained below [4][11]*

$$\boldsymbol{Numberofadoptersattimet = p(Remainingpotential) + q(Cumulativeadoptersatt - 1)}$$

$$\boldsymbol{(Remainingpotential)} \ \ldots\ldots\ldots\ldots \textbf{(Equation 1)}$$

*where,*

$$p(Remaining\ potential) - denotes\ the\ effect\ of\ innovation\ on\ adoption\ of\ the\ product$$

$$q(Cumulative\ adopters)(Remaining\ potential) - represents\ the\ imitation\ effect\ on\ adoption\ of\ the\ product$$

*The discrete mathematical form of the Bass model equation is shown below [4] [11].*

$$\boldsymbol{n(t) = p[m - N(t-1)] + \frac{q}{m}N(t-1)[m - N(t-1)]}$$

$$\boldsymbol{n(t) = pm + (q-p)N(t-1) - \frac{q}{m}N(t-1)^2} \ \ldots\ldots\ldots\ldots \textbf{(Equation 2)}$$

***where,***

$$n(t) - Number\ of\ adopters\ who\ will\ purchase\ the\ product\ at\ time\ t\ N(t-1) - Cumulative\ number\ of\ adopters\ of\ the\ product\ upto\ time\ period\ t-1$$

$$p - coefficient\ of\ innovation$$

$$q - coefficient\ of\ imitation$$

$$m - market\ potential$$

*Mathematical solutions to the Bass model are shown below [4][11].*

$$\boldsymbol{Cumulativenumberofadopters: \ N(t-1) = m\frac{[1-e^{-(p+q)(t-1)}]}{[1+\frac{q}{p}e^{-(p+q)(t-1)}]}} \ \ldots\ldots\ldots\ldots \textbf{(Equation 3)}$$

$$\boldsymbol{Numberofadoptersatagivetimet: \ n(t) = m\frac{p(p+q)^2[e^{-(p+q)t}]}{[p+qe^{-(p+q)t}]^2}} \ldots\ldots\ldots\ldots \textbf{(Equation 4)}$$

$$\boldsymbol{Timeofpeakadoptions: \ T^* = -\frac{1}{(p+q)}ln\frac{p}{q}} \ldots\ldots\ldots\ldots \textbf{(Equation 5)}$$

$$\boldsymbol{Numberofadoptersatthepeaktime: \ n(T^*) = \frac{1}{4q}(p+q)^2} \ldots\ldots\ldots\ldots \textbf{(Equation 6)}$$

The Bass model generates four key outputs: the cumulative number of adopters, the number of adopters at a specific time, the peak adoption time, and the number of adopters at peak time. These outputs can be determined using equations 3 to 5.

Figure 6 illustrates the cumulative sales curve, showing the number of adopters over a given period. The cumulative sales curve adopts an S-curve pattern, indicating that (i) the initial rate of sales growth is slow, (ii) sales growth accelerates significantly midway, and (iii) the growth rate eventually plateaus as market saturation is reached [7].

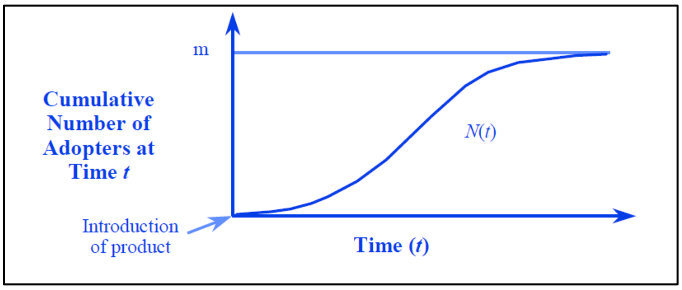


**Figure 6: Cumulative sales curve [11]**

The sales curve, depicted as a bell curve in Figure 7, represents the number of adopters at any given time (t). According to Figure 7, the number of adopters increases until time (T*), then gradually decreases as the market saturates. The sales curve identifies (T*), the peak adoption time, and n(T*), the peak number of adopters during the product life cycle.

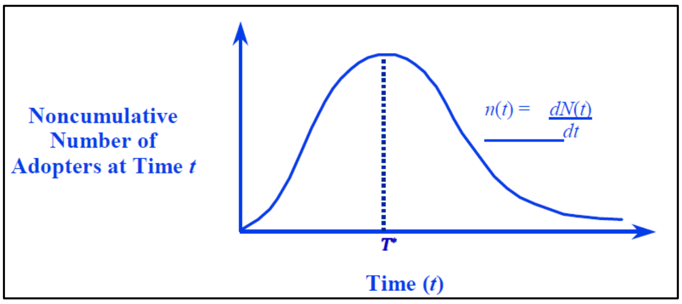


**Figure 7: Sales curve [11]**

The timing of peak product adoption is influenced by two key drivers: internal and external influences.

(a) Internal influences: When the coefficient of imitation is greater than the coefficient of innovation (q > p), internal influences drive product adoption. This suggests that word of mouth has a greater impact than curiosity on the adoption of the new product. The peak adoption rate occurs at a time (T*) after the product launch. Figure 8 illustrates the adoption curve shaped by internal factors.

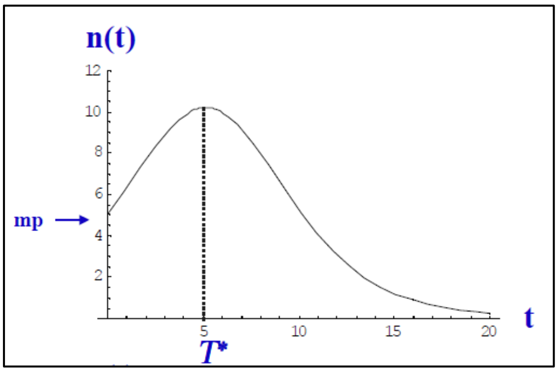


**Figure 8: Cumulative sales curve [11]**

(b) External influences: Conversely, when the coefficient of imitation is less than the coefficient of innovation (q < p), external factors drive adoption. This indicates that word of mouth has less impact than curiosity on the product's adoption. The sales rate declines after the launch. Figure 9 depicts an exponentially declining adoption curve influenced by external factors.

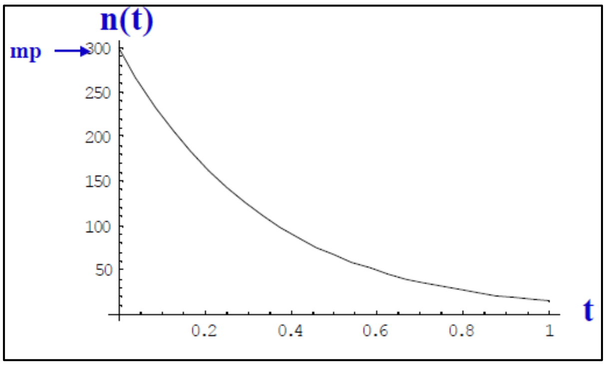


**Figure 9: Adoptions due to external influences [11]**

Figure 10 provides a reference for the innovation and imitation parameters of various products [11]. In all the listed products, the imitation parameter exceeds the innovation parameter, indicating that product adoptions are primarily driven by internal influences.

| Product/ Technology | Innovation parameter ($p$) | Imitation parameter ($q$) |
|---|---|---|
| B&W TV | 0.028 | 0.25 |
| Color TV | 0.005 | 0.84 |
| Air conditioners | 0.010 | 0.42 |
| Clothes dryers | 0.017 | 0.36 |
| Water softeners | 0.018 | 0.30 |
| Record players | 0.025 | 0.65 |
| Cellular telephones | 0.004 | 1.76 |
| Steam irons | 0.029 | 0.33 |
| Motels | 0.007 | 0.36 |
| McDonalds fast food | 0.018 | 0.54 |
| Hybrid corn | 0.039 | 1.01 |
| Electric blankets | 0.006 | 0.24 |

**Figure 10: Parameters for other product categories [11]**

## VI. IMPLEMENTATION

Figure 11 provides a flow chart summarizing the implementation steps for forecasting appliance demand.

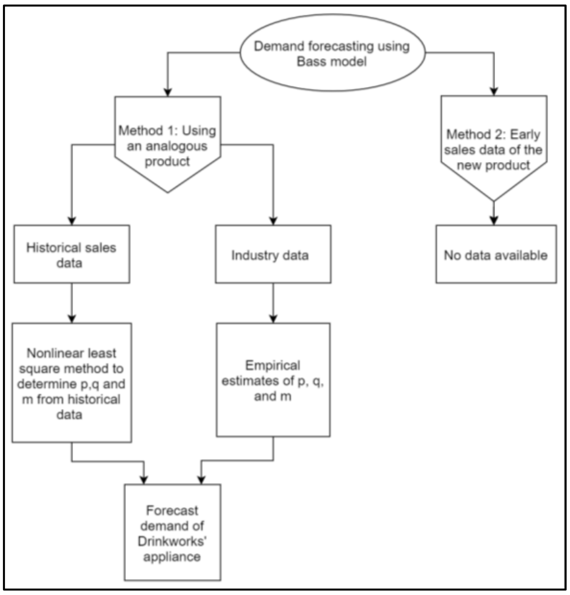


**Figure 11: Implementation process flow**

Method 1: Using an Analogous Product

The initial step in forecasting the appliance's demand involved estimating the three model parameters using the analogous product method (Method 1). Method 2 was excluded as there was no early sales data available, given the product had yet to be launched. Within the analogous product method, two distinct approaches were employed to estimate the model parameters [12]:
(a) Historical sales data
(b) Industry data

In this method, the initial step involved selecting an analogous product previously launched, whose data could serve as a reference. This was accomplished by identifying products with existing sales data and similar functions and features to Drinkworks' new product. Keurig's coffee appliance and a generic automated coffee maker were the two analogous products shortlisted for their comparable characteristics.

A comprehensive set of attributes that significantly influenced the adoption patterns of the new product was identified for comparison. Based on the product comparison method proposed by Robert J. Thomas (1985), the selected attributes encompassed diverse aspects of the product, including the environment, market structure and strategy, and product characteristics [13]. Table 1 compares the three analogous products in terms of their complexity, convenience, innovativeness, cost, market potential, regulations, and technology.

**Table 1: Comparison of analogous products**

| **Attributes** | **Analogous Products** | | | **Factor Weight** |
|---|---|---|---|---|
| | **Drinkworks' Appliance** | **Keurig's Coffee Appliance** | **Coffee Maker** | |
| Complexity | X | X | | 5 |
| Convenience | X | X | X | 5 |
| Innovativeness | X | X | X | 5 |
| Cost | X | X | X | 4 |
| Market potential | X | X | X | 4 |
| Regulations | X | | | 3 |
| Technology | X | X | | 5 |
| Total score | **31** | **28** | **18** | - |

An (X) was assigned to a product if it possessed a specific attribute. Factor weights were allocated to each attribute based on their influence on the product's adoption decision. The total attribute scores of the three products were calculated, leading to the conclusion that Keurig's coffee appliance was the most suitable analogous product, as its score closely matched that of Drinkworks' new appliance [12].

(a) Historical Sales Data

The next step in the parameter estimation process involved refining the historical sales data of the analogous product before estimating the model parameters for Drinkworks' new appliance. The available historical sales data represented the weekly number of coffee appliances sold to retailers by Keurig. Ideally, actual point-of-sales data would have been more appropriate, as it accurately reflects the product's adoption pattern. However, this data was inaccessible due to the unavailability of retailers' sales information.

Keurig's sales-to-retailers data was available for four different SKUs of coffee appliances, belonging to the same product family, and segregated based on retailers across the US. The historical sales data spanned approximately 2.5 years, available on a weekly basis. This data was pooled across the four SKUs and retailer locations to derive an accurate parameter estimate representing the national-level adoption of one product family.

In the final step, the refined data was used to estimate the model parameters p and q using the nonlinear least squares estimation method. This method minimizes the sum of squares of the defined function to estimate the model parameters. Nonlinear least squares analysis was performed on 128 weeks of sales-to-retailers data to evaluate the weekly parameter estimates. The parameters p and q, estimated from the analogous product's historical data, were assumed to apply to Drinkworks' new appliance. The market potential parameter m was externally estimated by Drinkworks' marketing team using market research techniques and expert opinions. Two different market potential parameters were evaluated for the pilot and US national launch phases, and it was decided to use these externally determined parameters, as they better represented the actual markets for Drinkworks' appliance.

Model parameter estimates depend on the historical data's time period. Therefore, the historical sales data's time period must align with the required forecasting time period to obtain accurate parameter estimates. Table 2 presents the model parameters estimated by aggregating historical sales data on a yearly, monthly, and weekly basis. Weekly and monthly parameter estimates were used to forecast demand on a weekly and monthly basis, respectively.

**Table 1: Parameters estimated using historical sales data**

| **Parameters** | **Yearly Data** | **Monthly Data** | **Weekly Data** |
|---|---|---|---|
| Coefficient of Innovation, p | 0.0522 | 0.0114 | 0.0035 |
| Coefficient of Imitation, q | 1.0622 | 0.12 | 0.0291 |
| Market Potential, m | Pilot: 600 and US national launch: 51000 | | |

(b) Industry Data

The second approach involved utilizing empirical parameter estimates of an analogous product derived from industry data. The work of Gary Lilien et al. on "Diffusion Models: Managerial Applications and Software" was referenced to leverage these empirical estimates [14]. However, the available empirical parameter estimates were limited to generic product categories. The "Coffee maker ADC" was identified as the most comparable product to Drinkworks' appliance among the available options. Consequently, the empirical parameter estimates of \(p\) and \(q\) for the "Coffee maker ADC" were directly applied as model parameters for Drinkworks' appliance, while the same externally evaluated market potential parameters for the pilot and US national launch were employed.

As previously mentioned, parameter estimates are contingent on the data's time period. Thus, three different empirical parameter estimates were derived by aggregating industry data across three different time periods. Table 3 presents the empirical estimates of model parameters based on yearly, monthly, and weekly industry data. Weekly and monthly empirical parameter estimates were utilized to forecast demand on a weekly and monthly basis, respectively.

**Table 3: Empirical parameter estimates [14]**

| Parameters | Yearly Data | Monthly Data | Weekly Data |
|---|---|---|---|
| Coefficient of Innovation, p | 0.077 | 0.0168 | 0.0052 |
| Coefficient of Imitation, q | 1.106 | 0.1250 | 0.0303 |
| Market Potential, m | Pilot: 600 and US national launch: 51000 | | |

Two distinct execution techniques were applied for the two parameter estimation approaches. For historical sales data, the nonlinear least squares estimation technique was utilized to estimate the model parameters. In contrast, for industry data, empirical parameter estimates were directly substituted into the output equations of the Bass model.

(a) Historical Sales Data

In this first approach, the model parameters were evaluated using nonlinear least square method from the available sale-to-the-retailers data of the analogous product (Keurig's coffee appliance).
From the equation 2,

$$n(t) = pm + (q-p)N(t-1) - \frac{q}{m}N(t-1)^2$$

Equation 2 was converted into a simple quadratic nonlinear equation,
$n(t) = a + bN(t-1) - cN(t-1)^2$ ............ (Equation 7)
The values of a, b, and c were evaluated using the nonlinear least square estimation method on the historical sales data of the analogous product. Once the values of a, b, and c were known, the model parameters were determined using equations 8 to 10. However, for this project the market potential parameter m was determined externally through market research and hence, equation 8 was used as a constraint on the values of a, b, and c while implementing the nonlinear least square estimation method. Subsequently, the parameters p and q were calculated using equations 9 and 10 [12].
$m = \frac{-b \pm \sqrt{b^2-4ac}}{2c}$ ........... (Equation 8)

$p = \frac{a}{m}$ ............ (Equation 9)

$q = p + b$ ............ (Equation 10)
After determining all the model parameters, the product life-cycle sales curve of Drinkworks' new appliance was generated using the output equations 3 to 6.

(b) Industry Data

In this second approach, empirically estimated model parameters (p) and (q) (for the product category "Coffee Maker ADC") were derived from industry data as detailed in the work of Gary Lilien et al. [14]. These empirical estimates for (p) and (q) were assumed to be applicable to Drinkworks' new appliance. The market potential parameter (m) was determined through market research techniques and expert opinions.

With all three model parameters known, the product life-cycle sales curve for Drinkworks' new appliance was calculated by inputting these parameters into output equations 3 to 6.

Python, a high-level programming language, was employed to code and execute the forecasts using the Bass model. A graphical user interface (GUI) was designed and integrated with the Python code, enabling users to interact with and swiftly execute the model. Figure 12 displays the home page of the Drinkworks DP tool, allowing users to select the product sub-category to be forecasted.


**Figure 12: DP home page**

Figure 13 shows the window that appears after selecting the appliance forecast option on the home page. Users can then choose one of the three models to forecast the appliance's demand.

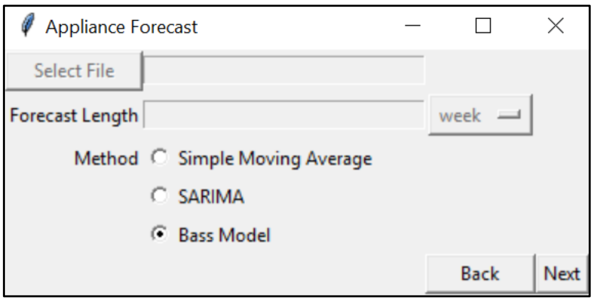


**Figure 13: Appliance forecast window**

Upon selecting the Bass model, users will be presented with the Bass model parameter window, as depicted in Figure 14. This page requires three main inputs: the forecast start date, the forecast length (in weeks or months), and one of the two parameter estimation approaches.

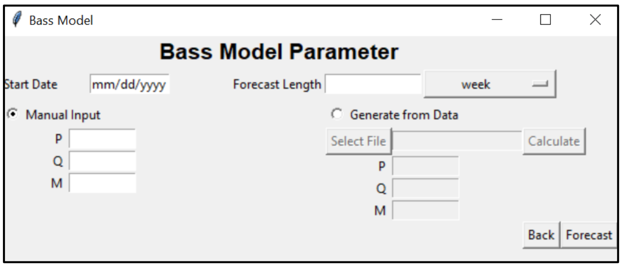


**Figure 14: Bass model parameter window**

Figures 15 and 16 depict the two distinct parameter estimation approaches employed to forecast the appliance demand during the 21-week pilot phase.

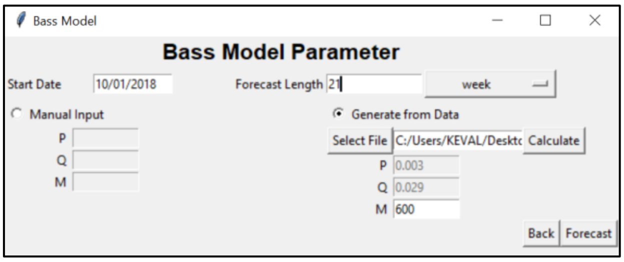


Figure 15: Parameter estimation using historical sales data

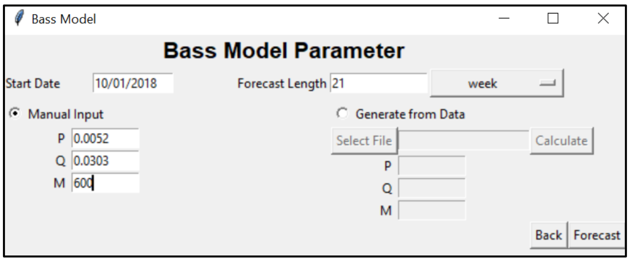


Figure 16: Empirical parameter estimates

Clicking on the forecast button generates the appliance demand forecast based on the provided input parameters. Figures 17 and 18 display the sales and cumulative sales curves for the pilot phase, produced by the DP tool using parameter estimates from historical sales data.

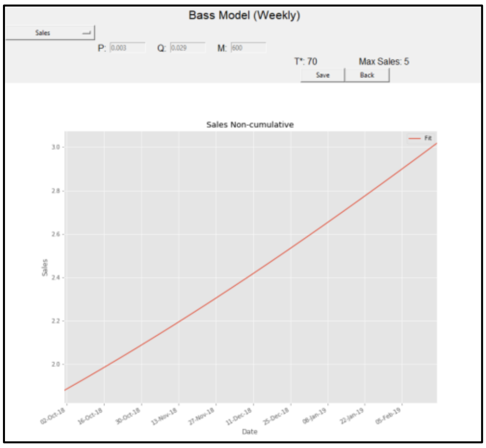


Figure 17: Sales curve results page

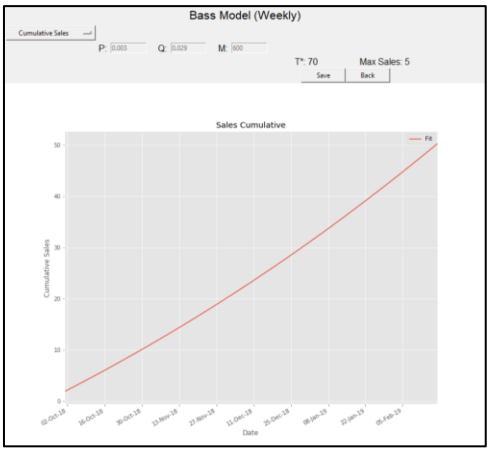


Figure 18: Cumulative sales curve results page

## VII. RESULTS

The DP tool, utilizing the Bass model, generated appliance forecasts for both the pilot and US national launch on a weekly and monthly basis.

Table 4 provides a summary of the appliance forecast results for the pilot, highlighting the outcomes of various parameter estimation approaches and time horizons.

**Table 4: Pilot forecast summary**

| Input data | Weekly forecast | | Monthly forecast | |
|---|---|---|---|---|
| | **Historical sales data** | **Industry data** | **Historical sales data** | **Industry data** |
| **Model Parameters** | p = 0.0031, q = 0.0291, m = 600 | p = 0.0052, q = 0.0303, m = 600 | p = 0.0114, q = 0.12, m = 600 | p = 0.0168, q = 0.125, m = 600 |
| **Model Results** | $T^*$ = 70 weeks and Maximum sales in a week = 5 appliances | $T^*$ = 49 weeks and Maximum sales in a week = 6 appliances | $T^*$ = 18 months and Maximum sales in a month = 22 appliances | $T^*$ = 14 months and Maximum sales in a month = 24 appliances |

Figures 19 to 22 present the sales forecasts of Drinkworks' new appliance for the pilot phase, based on various parameter estimation approaches and time horizons. All four figures demonstrate a linear increase in sales over the respective forecasting periods. In all cases where (q > p), this implies that the adoption of Drinkworks' new appliance is driven by internal factors. The model results in Table 4 show the time and magnitude of maximum sales anticipated for Drinkworks' new appliance.

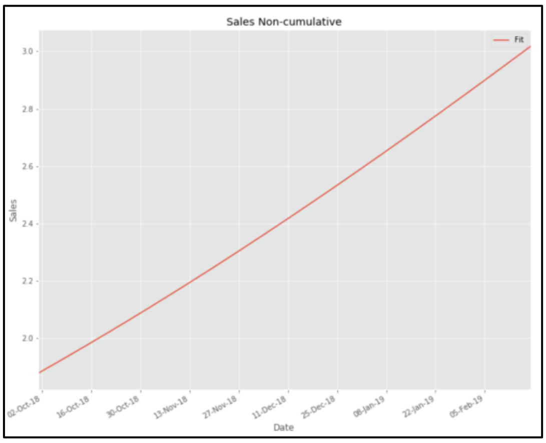


**Figure 19: 21-week forecast using historical sales data approach**

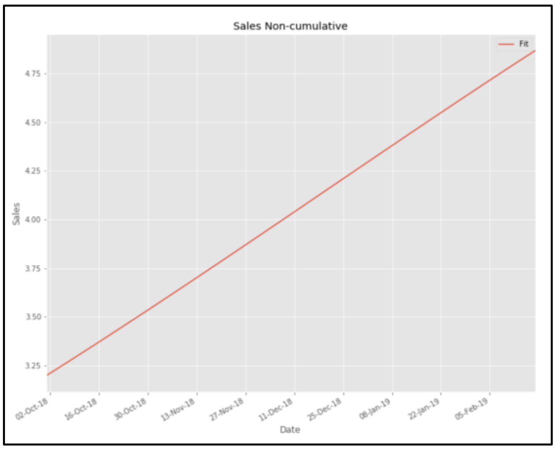


**Figure 20: 21-week forecast using industry data approach**

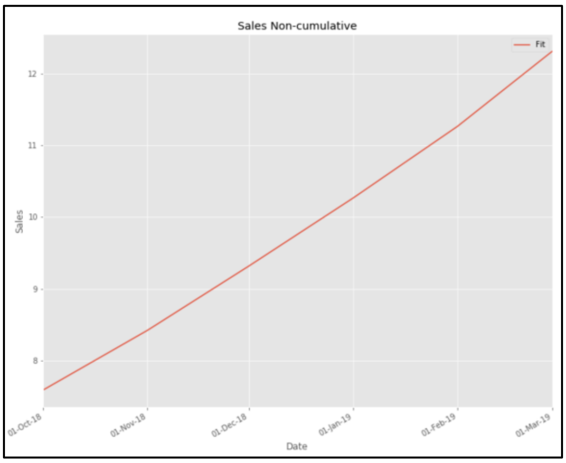


**Figure 21: 5-month forecast using historical sales data approach**

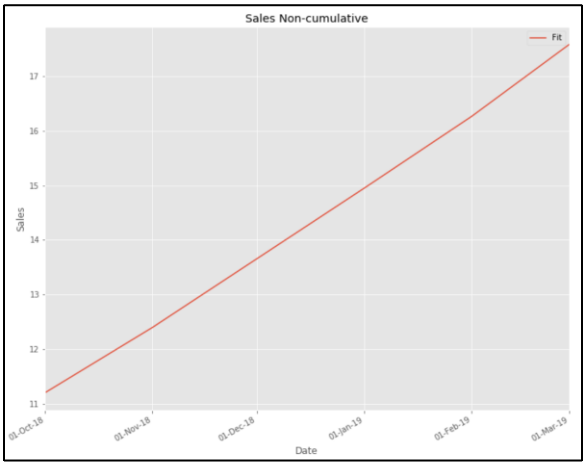


**Figure 22: 5-month forecast using industry data approach**

Figure 23 shows the monthly sales forecast for a 40-month pilot run (hypothetical scenario) using parameters estimated from historical sales data. Figure 24 presents the expected cumulative sales (S-curve) over the same 40-month pilot run (hypothetical scenario) based on the same parameter estimates.

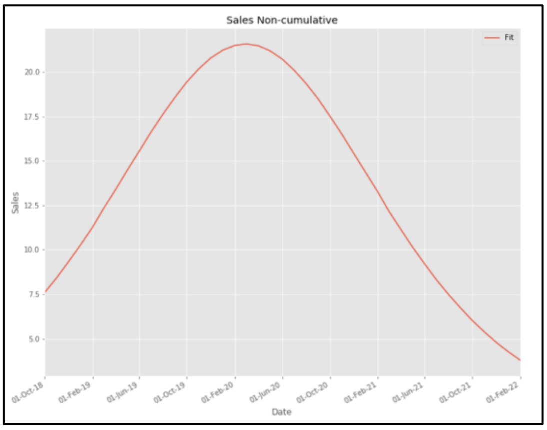


**Figure 23: 40-month pilot run sales forecast**

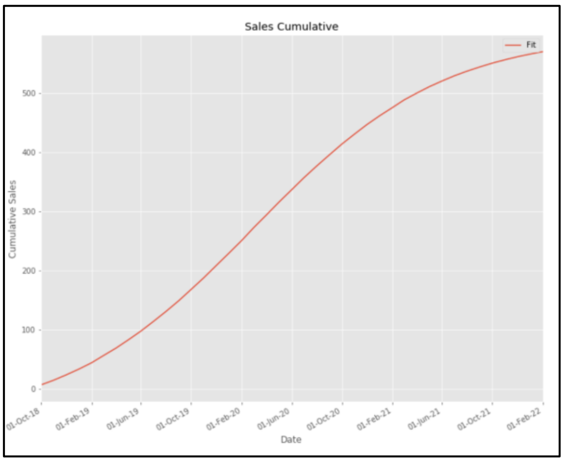


**Figure 24: Cumulative sales over 40 months of pilot run**

Table 5 provides a summary of the appliance forecast results for the US national launch, highlighting the outcomes derived from various parameter estimation approaches and time horizons.

**Table 5: US national launch forecast summary**

| **Input data** | **Weekly forecast** | | **Monthly forecast** | |
|---|---|---|---|---|
| | **Historical sales data** | **Industry data** | **Historical sales data** | **Industry data** |
| **Model Parameters** | p = 0.0031, q = 0.0291, m = 51000 | p = 0.0052, q = 0.0303, m = 51000 | p = 0.0114, q = 0.12, m = 51000 | p = 0.0168, q = 0.125, m = 51000 |
| **Model Results** | $T^*$ = 70 weeks and Maximum sales in a week = 452 appliances | $T^*$ = 50 weeks and Maximum sales in a week = 528 appliances | $T^*$ = 18 months and Maximum sales in a month = 1834 appliances | $T^*$ = 14 months and Maximum sales in a month = 2049 appliances |

Figures 25 to 28 present the sales forecasts of Drinkworks' new appliance for the US national launch, utilizing various parameter estimation approaches and time horizons. All four figures display a linear increase in sales. Since (q > p) in all scenarios, it can be inferred that the adoption of Drinkworks' new appliance is driven by internal influences. Table 5 indicates the expected time and magnitude of maximum sales during the US national launch.

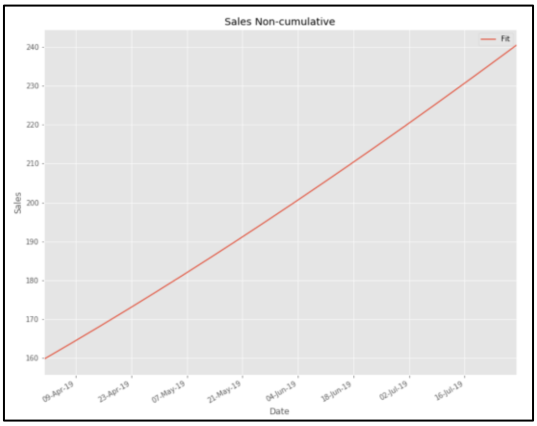


**Figure 25: 18-week forecast using historical sales data approach**

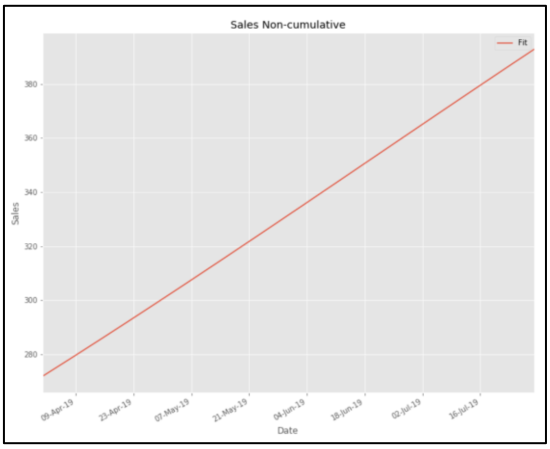


**Figure 26: 18-week forecast using industry data approach**

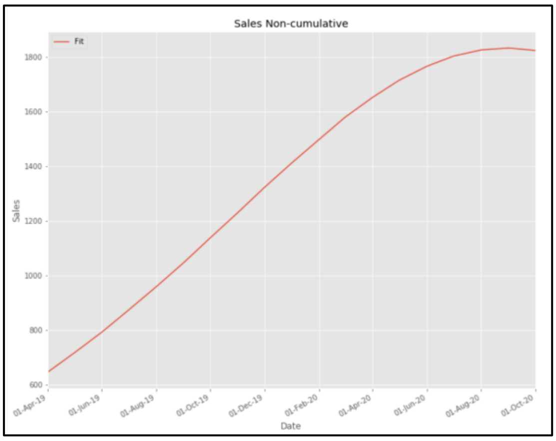


**Figure 27: 18-month forecast using historical sales data approach**

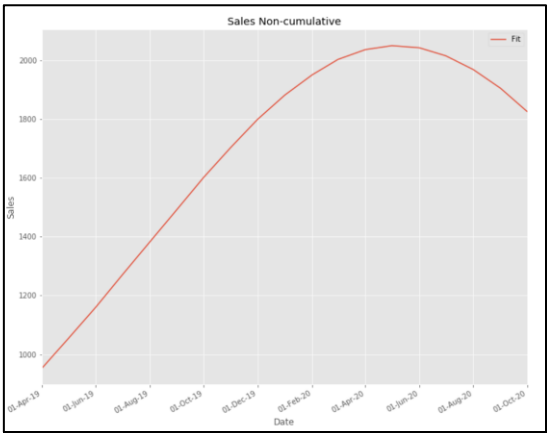


**Figure 28: 18-month forecast using industry data approach**

Figure 29 illustrates the monthly sales forecast over a 40-month period based on parameters estimated from historical sales data. Figure 30 shows the expected cumulative sales (S-curve) during the 40 months of the US national launch, using the same parameter estimates.

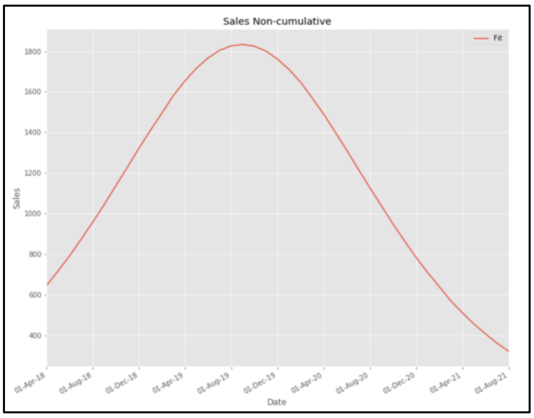


**Figure 29: 40-month US national launch sales forecast**

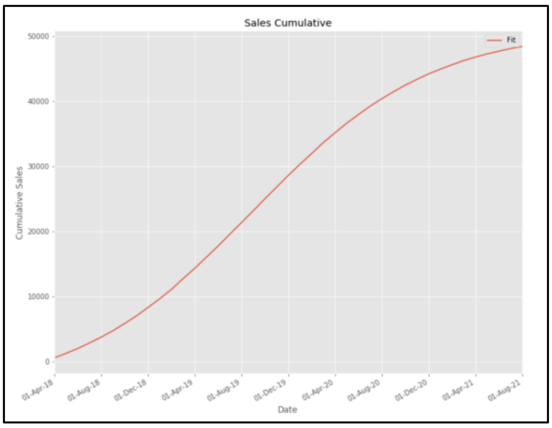


**Figure 30: Cumulative sales over 40 months of US national launch**

Comparing the forecasting results in Tables 4 and 5 reveals that the two parameter estimation approaches predicted different Bass model outputs. Such varied results were anticipated due to differences in the input model parameters and the analogous products used.

Ultimately, the decision was made to use the forecast results generated by the first parameter estimation approach (using historical sales data of an analogous product—Keurig's coffee appliance). This decision was based on two key rationales:

(a) As shown in Table 1, Keurig's coffee appliance had a higher score tally, indicating it was more analogous to Drinkworks' new appliance.

(b) Unlike the empirical data, the historical sales data of the analogous product was well-established and verified.

Thus, estimating model parameters from Keurig's coffee appliance was deemed the best approach for producing a more accurate forecast for Drinkworks' new appliance. However, after completing the pilot, it is recommended to use Method 2 (i.e., the pilot's sales data) to estimate the model parameters for the US national launch. This approach is expected to result in a more accurate estimation of the model parameters for Drinkworks' new appliance. Figure 31 summarizes the final appliance forecasting approach using the Bass model.

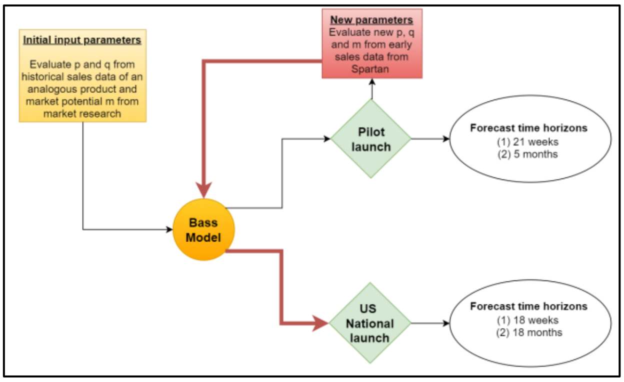


**Figure 31: Final forecasting approach**

## VIII. PODS FORECASTING

Pods are single-use disposable units containing alcohol and a flavoring agent necessary to prepare specific drinks. The consumption of these pods is dependent on the number of appliances sold. Additionally, the rate of pod consumption varies according to consumers' drinking behaviors. Therefore, it is crucial to consider all these factors when forecasting pod consumption. Ultimately, the pod consumption forecast will serve as the baseline for demand prediction, taking into account the dynamics of the supply chain system.

All Drinkworks appliances were equipped with a "connect kit," enabling real-time data collection and transmission via internet and Bluetooth connectivity. These appliances can record pod consumption data by SKU, date, time of consumption, and appliance serial number. The collected data is transferred to Drinkworks' online database for analyzing consumer behavior and enhancing forecasting efforts. However, data transmission is contingent upon customer approval.

To test the "connect kit" program, Drinkworks distributed 42 appliances to employees before the pilot and national launch in the US. Of these, only 28 appliances were actively used and transmitting consumption data. The data from these 28 appliances were utilized to evaluate the average consumption rate and demonstrate the pod forecasting approach. The extracted consumption data were aggregated weekly based on appliance serial numbers.

The subsequent step involved segregating and grouping consumers based on their consumption rates into two clusters: (i) infrequent users and (ii) frequent users. K-means cluster analysis was chosen to identify

groups from the unlabeled data. This analysis was performed on the available consumption data using Minitab software. Eleven weeks of consumption data from the 28 appliances were input into the software, which was instructed to divide the data into two clusters. The software algorithm employed a two-step iterative process to produce the final results:
(a) Data assignment – Each cluster was represented by one centroid (the arithmetic mean of data points), and each data point was assigned to the nearest centroid based on squared Euclidean distance. (b) Centroid update – The centroids were recalculated by averaging all data points assigned to each cluster's centroid.
The algorithm continually iterated these two steps until no data point was reassigned to a different cluster [15]. Figure 5.1 illustrates the clustered consumption data from the 28 appliances, aggregated on a weekly basis. Cluster A represents infrequent users, while Cluster B represents frequent users.

| Count of pods consumed | Appliance Serial Numbers | | | | | | | | | | | | | | | | | | | | | | | | | | | | Total pods |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 4 | 5 | 7 | 8 | 10 | 12 | 13 | 15 | 18 | 19 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 30 | 31 | 32 | 33 | 34 | 36 | 38 | 39 | 40 | 42 | |
| 3/19/2018 - 3/25/2018 | 7 | 5 | - | - | 8 | - | 1 | - | - | - | 23 | - | - | 7 | - | - | 3 | 1 | 5 | 10 | 7 | 3 | 21 | - | 1 | 3 | 8 | - | 113 |
| 3/26/2018 - 4/1/2018 | 7 | 6 | | 6 | 27 | 10 | 12 | 7 | 4 | - | 9 | 2 | 9 | 12 | 21 | 18 | 11 | 4 | 24 | 22 | 4 | 6 | 3 | - | 4 | 7 | 4 | - | 239 |
| 4/2/2018 - 4/8/2018 | - | 8 | 5 | - | 10 | 10 | 3 | - | 2 | 4 | 2 | 3 | 2 | 9 | 33 | 3 | 5 | 1 | 4 | 9 | 4 | 2 | 11 | - | 4 | 1 | 2 | 16 | 153 |
| 4/9/2018 - 4/15/2018 | - | - | 7 | - | - | 6 | - | 2 | - | 9 | - | - | - | - | 3 | 4 | 1 | 3 | 3 | 12 | 2 | 2 | 6 | - | - | 8 | - | 20 | 88 |
| 4/16/2018 - 4/22/2018 | - | 2 | - | - | 24 | 4 | - | 3 | 1 | 12 | 2 | - | - | 2 | 13 | 2 | - | 2 | - | - | 3 | 6 | 1 | - | - | - | 6 | 5 | 88 |
| 4/23/2018 - 4/29/2018 | - | 3 | 3 | - | 17 | 1 | 1 | 1 | - | 1 | 2 | 1 | 7 | 11 | 21 | - | 1 | 2 | - | - | 2 | 4 | 15 | - | - | 4 | 1 | 6 | 104 |
| 4/30/2018 - 5/6/2018 | - | 6 | - | - | 5 | 5 | - | 4 | - | 3 | 3 | 2 | | 2 | 17 | - | - | - | 1 | - | - | 4 | - | - | - | 2 | 2 | 6 | 62 |
| 5/7/2018 - 5/13/2018 | - | 2 | - | - | 10 | 5 | - | 3 | 2 | - | - | - | 2 | - | 19 | - | 3 | - | - | - | 1 | - | 2 | 1 | - | - | 1 | 5 | 56 |
| 5/14/2018 - 5/20/2018 | - | | - | - | 4 | 1 | - | - | - | - | - | - | - | - | - | 2 | 1 | 2 | - | - | - | 4 | - | - | - | - | - | 14 | 28 |
| 5/21/2018 - 5/27/2018 | - | 5 | - | - | 16 | 5 | - | - | - | - | - | - | 7 | 2 | - | - | - | - | - | - | - | 11 | 3 | - | - | 6 | - | 24 | 79 |
| 5/28/2018 - 6/3/2018 | - | 5 | - | - | 1 | - | - | - | - | - | - | 2 | - | 1 | 1 | 4 | 1 | 1 | 2 | - | 6 | 3 | 1 | - | - | 1 | 3 | 25 | 57 |
| Cluster | A | A | A | A | B | A | A | A | A | A | A | A | A | A | B | A | A | A | A | A | A | A | A | A | A | A | A | B | |
| Total pods consumed per appliance | 14 | 42 | 15 | 6 | 122 | 47 | 17 | 20 | 9 | 29 | 41 | 10 | 27 | 46 | 128 | 33 | 26 | 16 | 39 | 53 | 29 | 45 | 63 | 1 | 9 | 32 | 27 | 121 | 1067 |

**Figure 32: Clustered consumption data**

The results of the cluster analysis revealed two distinct groups based on the average pod consumption rate, as presented in Table 6.

**Table 6: Cluster analysis results**

| **Clusters** | **Infrequent User Cluster** | **Frequent User Cluster** |
|---|---|---|
| **% of total users** | 89.29% | 10.71% |
| **Total users** | 25 | 3 |
| **Average pod consumption per appliance per week** | 3 | 11 |
| **Weighted average pod consumption per appliance per week** | 3.46 | |

The final step involved integrating inputs from the cluster analysis, cumulative appliances sold to retailers, percentage of active appliances, and appliance forecast to derive the pod consumption forecast. The mathematical formula used to forecast weekly pod consumption is provided as Equation 11 below:

$$\begin{aligned} &Weekly\ pod\ consumption\ forecast \\ &\qquad = (Cumulative\ appliances\ sold - to - the - retailers \\ &\qquad + Weekly\ appliance\ forecast) \\ &(\%\ Active\ appliances\ )(Weighted\ pod\ consumption\ per\ appliance\ per\ week\ ) \end{aligned}$$ ……….

(Equation 11)

$$\begin{aligned} &Weighted\ average\ pod\ consumption\ per\ appliance\ per\ week \\ &\qquad = \left(\frac{(Infrequent\ users\ \%)}{(Average\ consumption\ of\ light\ users\ per\ appliance\ per\ week)}\right) \\ &\qquad + \left(\frac{(Frequent\ users\ \%)}{(Average\ consumption\ of\ heavy\ users\ per\ appliance\ per\ week)}\right) \end{aligned}$$

…………(Equation 12)

As demonstrated in Equation 11, the formula requires four inputs to forecast the weekly pod consumption:
(a) Cumulative appliances sold-to-the-retailers: The actual cumulative number of appliances sold to retailers until the previous week. Appliance sales data to retailers was used instead of the actual retail sales data, as the latter was inaccessible.
(b) Weekly appliance forecast: The demand forecast for the appliance generated by the DP tool for the upcoming 'N' weeks.
(c) % Active appliances: The percentage of total appliances sold that are assumed to be in regular use. This parameter accounts for attrition in appliance usage and is assessed based on market research.
(d) Weighted average pod consumption per appliance per week: The weighted average pod consumption per appliance per week of both frequent and infrequent users. This can be calculated using Equation 12.
By substituting these parameters into Equation 11, the weekly pod forecast for the subsequent 'N' weeks can be determined.

Figure 33 displays the homepage of the Drinkworks DP tool, which enables users to select the product sub-category they wish to forecast.


**Figure 33: DP home page**

By selecting the pod forecast button, users are directed to the input parameter window of the pod forecast model. As shown in Figures 34 and 35, the model parameters are entered, and the software generates the pod consumption forecast using the mathematical formula described in Equation 11.

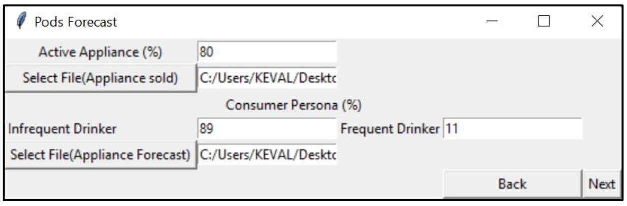


**Figure 34: Pod forecast parameter window**

Figure 35 displays the final pod forecasting results page of the DP tool.

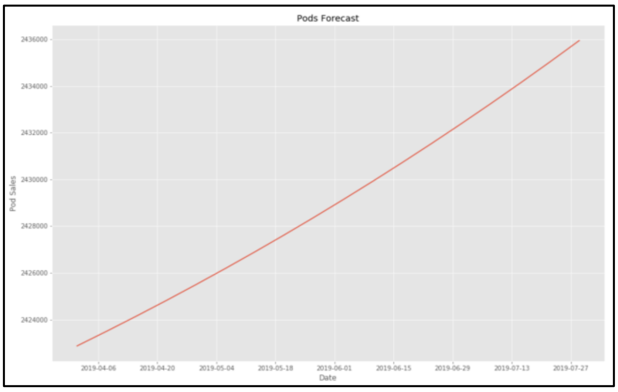


**Figure 35: Final results page**

The following section demonstrates the application of the pod forecasting model to predict pod consumption for the first 18 weeks of the US national launch. The input data used to generate the pod forecast were as follows:

(a) Cumulative appliances sold-to-the-retailers: Dummy data (Keurig's coffee appliance sold-to-the-retailers) was utilized to demonstrate the pod forecasting model, as Drinkworks' appliance had not yet been launched.

(b) Weekly appliance forecast: The appliance forecast generated by the Bass model for the first 18 weeks of the US national launch served as the input.

(c) % Active appliances: Based on market research, it was assumed that 80% of the total appliances would be active.

(d) Weighted average pod consumption per appliance per week: As calculated in Table 6, the consumption data from 28 active appliances was used to evaluate the weighted average pod consumption rate using Equation 12.

By substituting these input parameters into Equation 11, the model generated an 18-week pod consumption forecast. Figure 36 illustrates the pod consumption forecast obtained from the DP tool. This pod forecast serves as a demonstration (not the actual forecast) of the application of the pod forecasting model used by the DP tool.

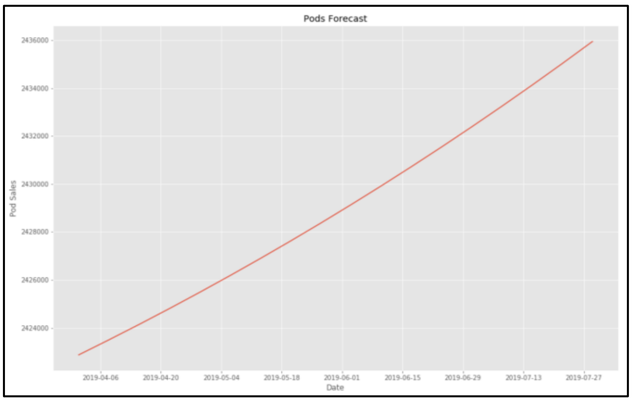


**Figure 36: Pod forecasting results**

## IX. Conclusion

The successful completion of this thesis project has provided Drinkworks with a comprehensive demand planning software tool for forecasting demand for their new product, the “in-home alcohol drink system.”

The final product of this thesis project is a tool that Drinkworks can use to forecast demand for their new product utilizing four different forecasting models (three appliance forecasting models and one pod forecasting model). This variety of models allows Drinkworks to forecast demand based on product sub-categories, the product life-cycle stage, and available historical data. The DP tool enables Drinkworks to generate strategic, tactical, and operational level forecasts, which can be used as inputs for material resource planning, production planning, and sales and operations planning.

The Bass model was utilized to generate a strategic forecast for the appliance launch. During the early launch phase, the simple moving average model will be used for operational forecasts. After collecting one and a half years of sales data, the SARIMA model will be employed to forecast demand, based on accumulated sales data. MAPE (Minimum Absolute Percentage Error) will be calculated weekly on a national level to evaluate forecast accuracy. Drinkworks will set an initial weekly MAPE target of 30%, based on Keurig’s experience, with the benchmark subject to change as market awareness increases.

The DP software tool will be integrated with Drinkworks’ online database for remote access by relevant parties. The DP tool can exchange input and output data in CSV format, a ubiquitous data file format compatible with various data analysis tools. The graphical user interface of the DP tool allows users to interact with and review forecasting results.

In summary, the DP tool will assist Drinkworks in systematically forecasting demand for their new product, the “in-home alcohol drink system,” throughout different product life-cycle phases and with varying levels of historical data. The DP tool will also be used to forecast demand for all future new products planned for launch.

## References


[1] P. Sun, "New Product Forecasting of Appliance and Consumables: Moving Average Model," Cambridge, 2018.

[2] Y. Li, "New Product Forecasting of Appliance and Consumables: SARIMA Model," Cambridge, 2018.

[3] J. C. Chambers, S. K. Mullick and D. D. Smith, "How to Choose the Right Forecasting Technique," *Harvard Business Review*, July 1971.

[4] F. M. Bass, "A New Product Growth Model for Consumer Durables," *Management Science* , vol. 50, no. 12 , pp. 1825-1832, 2004.

[5] R. Nau, "Statistical forecasting," August 2014. [Online]. Available: https://people.duke.edu/~rnau/Notes_on_forecasting_with_moving_averages--Robert_Nau.pdf. [Accessed June 2018].

[6] D. A. Anzalone, "Forecasting Fundamentals," Hillsborough Community College, Plant city, 2010.

[7] S. Radas, "Diffusion Models in Marketing: How to Incorporate the Effect of External Influence?," 2005. [Online]. Available: https://hrcak.srce.hr/file/28748. [Accessed May 2018].

[8] W. W. LaMorte, "Diffusion of Innovation Theory," Boston University School of Public Health, Boston , 2016.

[9] D. Jukić, "Total least squares fitting Bass diffusion model," *Elsevier* , vol. 53, no. 9-10, pp. 1756-1770, 2011.

[10] F. M. Bass, T. V. Krishnan and D. Jain, "Why the Bass model fits without decision variables," *Marketing Science,* vol. 13, no. 3, 1994.

[11] G. E. Fruchter, "Forecasting the Sales of New Products and the Bass Model," Bar-Ilan University, Tel Aviv, 2000.

[12] F. Ganjeizadeh, H. Lei, P. Goraya and E. Olivar, "Applying Looks-LIke Analysis and Bass Diffusion Model Techniques to Forecast a Neurostimulator Device with No Historical Data," *Elsevier*, no. 11, pp. 1916-1924, 2017.

[13] R. J. Thomas, "Estimating Market Growth for New Products: An Analogical Diffusion Model Approach," *Journal of Product Innovation Management*, no. 2, pp. 45-55, 1985.

[14] G. Lilien, A. Rangaswamy and C. V. d. Bulte, "Diffusion Models: Managerial Applications and Software," Institute of the Study of Business Markets, Philadelphia, 1999.

[15] A. Trevino, "Introduction to K-means Clustering," Oracle Data Science, 2016.