Introdução

As organizações passam atualmente por um momento de alta criticidade sobre seus processos onde a necessidade de assertividade sobre as medições de seus processos impactam diretamente sobre os custos e consequentemente o seu posicionamento no mercado global (Teoh, 2018). A globalização dos mercados e o aumento da competitividade entre as cadeias de suprimentos fizeram com que as decisões tomadas em cada elo tenham efeito sobre todas as organizações ligadas, desta forma, conforme Porter (2010) as organizações mais competitivas não são as que possuem os processos mais competitivos, mas os que estão inseridos na cadeia de suprimentos mais competitiva.

As previsões de demanda realizadas sobre cada elo da cadeia de suprimentos refletem sobre toda a cadeia impactando sobre diversos aspectos, sejam custos, performance e qualidade atendida. Uma previsão de demanda assertiva reflete de forma positiva sobre estes fatores evitando desperdícios ou baixo nível de serviço, contudo, segundo Gianesi e Biazzi (2011) embora as previsões de demanda sejam de extrema importância para os resultados das organizações e seus parceiros em toda a cadeia de suprimentos ainda tem sido tratada de modo pouco formal pelas organizações gerando assim desperdícios.

O volume de dados gerados nas cadeias produtivas cresce anualmente de forma exponencial, contudo, a utilização destes dados ainda é relativamente baixa. Um dos fatores que contribui sobre este fator é a baixa utilização de métodos formais na análise de dados e na tomada de decisões, estas por vezes realizada por métodos subjetivos advindos de experiências ou características do momento.

Neste sentido o presente trabalho busca a exploração para aplicação de métodos de previsão de demanda aplicados em uma indústria alimentícia de sorvetes situada na região metropolitana de Porto Alegre-Brasil, onde, por ser um produto cuja demanda é sensível às variações de temperatura, espera-se estimar com maior assertividade a demanda futura com base na temperatura prevista para os próximos períodos.

Como método matemático para modelagem e estimativa foram aplicados Interpolação de Lagrange e Regressão Linear para as variáveis históricas de demanda e temperatura, posteriormente os modelos foram comparados sobre os resultados apresentados a fim de definir o melhor método a ser aplicado ao problema específico.

Desenvolvimento

2.1 Referencial Bibliográfico

A realização de um processo de previsão de demanda necessita de uma série de etapas de coleta dos dados, tratamento e análise com o objetivo assegurar uma maior precisão dos valores estimados. Existem diversos métodos e ferramentas que podem ser aplicados pelas organizações com o objetivo de estimar os valores futuros, contudo, estas devem avaliar a efetividade da aplicação destes métodos conforme cada processo estimado (CORRÊA, GIANESI, CAON, et al. 2013).

2.2 Previsão de Demanda

Os métodos de previsão são utilizados de forma massiva pelas organizações com o objetivo de prever o comportamento dos seus processos e reduzir os desperdícios entre as etapas. Para Moreira (2001) os métodos de previsão de demanda podem ser classificados entre quantitativos e qualitativos onde se diferenciam pelo tipo de instrumento utilizado e os conceitos que formam a previsão. O autor cita que os métodos quantitativos se dividem em métodos causais que se caracterizam por modelos de regressão e métodos de séries temporais aplicadas a dados históricos.

O presente trabalho aplicou modelos matemáticos para regressão dos dados quantitativos de demanda diária para prever o futuro baseado na influência das variações de temperatura nos locais de venda, este método segundo GAITHER e FRAZIER (2001) exigem informações quantitativas preliminares para que possam ser modelados e posteriormente avaliados.

O método de Interpolação de Lagrange permite a modelagem de sistemas não-lineares, se aplicados a três pontos ou mais, de forma a aproximar os valores estimados entre os pontos de referência gerando o polinômio Pn(x) de grau =< n que interpola f em x0, ..., xn onde podemos representar Pn(xi) = y0L0(x) + y1L1(x) + ... + ynLn(x), onde os polinômios Lk(x) são de grau n (Ruggiero e Lopes; 1996).

Modelagem do Problema

A aquisição dos dados foi realizada através do software BW, ferramenta de *Business Inteligence* vinculado ao sistema ERP SAP onde foram extraídos os dados de demanda diária, em caixas, da unidade de negócios para o período de dois anos entre janeiro/2014 e dezembro/2015. Estes dados foram exportados em arquivo .CSV e importados no *software R,* utilizado no processo de tratamento dos dados e modelagem da solução.

Os dados históricos de demanda foram conciliados com as informações de temperatura diária extraídos do Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) apresentados no Instituto Nacional de Meteorologia (INMET). A informação de temperatura utilizada foi a média registrada na data da venda dos produtos, de forma a refletir as condições que geraram o volume de venda do período.

Para análise da Normalidade dos dados históricos foi aplicado o teste de Shapiro Wilk onde o resultado do P-Value foi de XXX onde pode-se concluir que não há normalidade. Como segundo XXXX os dados aplicados em métodos de regressão e interpolação necessitam apresentar normalidade em suas variações, foi necessário a limpeza dos valores.

Para identificação de *Outliers*, ou valores fora do padrão foram gerados gráficos de *Boxplot* conforme a figura 2 onde os valores acima da amplitude foram removidos de forma não influenciarem sobre os modelos treinados para estimar a demanda futura. Estes valores extraídos representam demandas fora da normalidade, onde outros fatores são preponderantes para estes valores.

Para análise dos dados foi realizado o teste de normalidade de Shapiro Wilk onde o resultado apresentado rejeitou a hipótese nula com um *P-Value* de 0.10, portanto, os dados apresentaram normalidade em sua distribuição e foi possível a sequência da aplicação dos algoritmos de previsão.

Para aplicação posterior da validação dos métodos e análise dos erros, os registros foram separados pelo método de *Hold-out validation*, gerando 80% dos dados para treino e 20% dos dados para teste/validação dos modelos.

Os registros de demanda foram agrupados conforme a temperatura gerando uma tabela média de demanda, tabela 1, consolidada por temperatura no dia gerando a correlação xi,yi para as demandas e temperaturas.

O método de interpolação foi aplicado para os cinco pontos com maior quantidade de registro de demanda (n) que variaram entre 20°C e 25°C, e para os pontos fora deste intervalo foi aplicado o método de ajuste de curvas. O modelo de interpolação gerado para estes registros foi aplicado aos dados segregados para teste, estimando para cada temperatura o valor interpolado.

O objetivo da pesquisa é estimar a demanda de vendas prevista com base na temperatura registrada no período através do levantamento de dados históricos, análise de sazonalidade e aplicação de inferências estatísticas onde diversos são os fatores que interferem nas demandas de vendas nas organizações e podem auxiliar na previsão futura das demandas. A assertividade sobre as previsões de demanda impacta sobre toda a cadeia produtiva na redução das necessidades de estoque e produção para atendimento apenas da demanda específica do cliente, desta forma, possibilitam redução de custos em escala sobre toda a cadeia produtiva. Foram extraídos dados de um ano de venda diária em uma indústria de sorvetes devido a sensibilidade da demanda deste segmento pelas variações de temperatura. Utilizando a Linguagem de Programação Estatística R, foi realizado sobre estes dados o processo de limpeza dos dados identificando o período de sazonalidade onde foram removidos os seis meses de menor demanda entre abril e setembro e dias úteis, visto que, os faturamentos aos domingos ou feriados são esporádicos e não representativos ao modelo. Para identificação dos *outliers* nos dados foi utilizada a função *boxplot* que apresentou dados acima da amplitude, estes foram removidos para que os mesmos não interferissem sobre os algoritmos de previsão e garantissem a normalidade dos dados. Para análise dos dados foi realizado o teste de normalidade de Shapiro Wilk onde o resultado apresentado rejeitou a hipótese nula com um *P-Value* de 0.10, portanto, os dados apresentaram normalidade em sua distribuição e foi possível a sequência da aplicação dos algoritmos de previsão. Os dados foram separados entre treino e teste na proporção de 80% de treino e 20% para teste, o agrupamento das médias foi realizado sobre os dados de treino, consolidando as temperaturas a cada 1ºC para utilização destes dados como referência nos algoritmos de previsão. Devido aos registros estarem concentrados entre os registros de 16º e 25ºC, foi aplicado o método de Interpolação sobre os dados neste intervalo e para os registros acima e abaixo destas temperaturas, o método utilizado foi de Ajuste de Curvas. Os dados utilizados para esta comparação foram os registros separados para Teste onde os resultados estimados foram cruzados com os registros efetivos do modelo e aplicado o cálculo do Erro Absoluto e Erro Relativo, sendo este último com um resultado médio de 30%. Também foi possível observar que os resultados estimados estiveram em todos os registros entre o 1º e 3º quartil. Nos resultados foi possível observar que resultados estimados pelos métodos de interpolação e ajuste de curvas apresentaram para um mesmo registro de temperatura um erro 2% e de 40%, fato que pode ser explicado por outras variáveis que impactam sobre a demanda que precisam ser aplicados ao modelo para aumento da assertividade, estes temas de sequência do projeto através da aplicação de algoritmos de *Machine Learning*.