|  |  |  |  |
| --- | --- | --- | --- |
| Análise RFV | | | |
| **Data** | **Cientista Responsável** | **Grupo Responsável** | **Consultoria Envolvida?** |
| 24/07/2023 | Diullio Pereira Dos Santos | Planejamento | NÃO |

Sumário

[1 Objetivos do Projeto 1](#_Toc130203927)

[2 Requisitos Mínimos 1](#_Toc130203928)

[3 Escopo do Entregável 1](#_Toc130203929)

[4 Metodologia 1](#_Toc130203930)

[5 Premissas 1](#_Toc130203931)

[6 Variáveis utilizadas 2](#_Toc130203932)

[7 Fontes utilizadas 2](#_Toc130203933)

[8 Dicionário de dados 2](#_Toc130203934)

[9 Cronograma de execução 2](#_Toc130203935)

# Objetivos do Projeto

Análise de Recência, Frequência e valor que resulta em agrupamento dos pdv’s para entendimento dos mais valiosos.

# Requisitos Mínimos

* Uso da base de venda sell-in (SQL).
* Python
* PostgreSQL
* PowerBI

# Escopo do Entregável

As relações entre os prestadores de serviços e as clientes têm se modificado pelo aumento da concorrência. O estudo de caso abordado teve como objetivo a análise da fidelização de clientes da Hypera Pharma, com o auxílio da metodologia RFV (Recência, Frequência e Valor).

O RFV é uma técnica usada para priorizar os clientes. Recência (R) nos diz quando é a última data de compra. Frequência (F) nos diz com que frequência eles fazem compras. Valor (V) nos diz quão grande é o tamanho médio do ticket. Na verdade, é o valor vitalício do cliente.

Por meio da pesquisa descritiva e quantitativa a partir do uso de dados, buscou-se identificar os clientes de maior valor para a empresa por serem fiéis e representativos quanto ao faturamento. Os dados tratados compreenderam uma leitura do comportamento de 761 clientes ao longo de 5 anos. Diante das análises realizadas, e da estratificação da base de clientes foi possível identificar comportamentos e reportar medidas de retenção aos clientes que não retornaram, bem como, direcionar novas estratégias para favorecer a fidelização dos clientes que frequentam a empresa com maior assiduidade.

Além disto foi criado um sistema de histórico, no qual pode ser observado a influência das ações em cada cliente, avaliando se ocorre a mudança de score e categoria do cliente. Por fim foi validado a metodologia via *machine learning* com aplicação do algoritmo de *RandomForest*.

# Metodologia

Para cálculo do score RFV, foi identificado o cliente através da rede, a recência através do último dia comprado, frequência através do número total de transações, e valor, sendo tamanho total de compras dividido pelo número total de transações.

Para definir a pontuação RFV em uma escala de 1 a 5 para cada cliente em termos de atualidade, frequência e valor monetário. 1 é o mais baixo e 5 é o mais alto. Foram classificados os registros frequência e valor procedendo a distribuição por quintis, estabelecendo notas de 1 a 5 para cada variável. Para a variável recência, adotou-se uma metodologia interna em parâmetros que seria interessante para equipe de inteligência de mercado buscando uma maior distinção dos períodos de compras dos clientes. Os scores foram definidos como 5 para até 2 meses da última compra, score 4 de 2 a 7 meses, 3 de 8 a 15 meses, 2 de 16 a 26 meses, e 1 para períodos acima. Após atribuída a nota pela perspectiva de análise de cada variável foi efetuada a concatenação das variáveis.

Foi criando um banco de dados SQL para registro histórico, podendo avaliar o desempenho das ações sobre os clientes com o decorrer do tempo, além de observar as mudanças de score e segmentos.

Para validação da metodologia, foi adotado um modelo de clusterização *k-means*, no qual o dataset continha os dados de recência, frequência, valor e classificação (segmentos). Através do *GridSearch* da biblioteca *sci-kit learn*. O método para *machine learn* utilizado foi o *RandomForestClassifier*. Para validação dos resultados foi gerado a matrix de confusão e por fim analisado os resultados estatísticos do método, como acurácia, precisão, revocação e *F1-score*.

# Premissas

O embasamento teórico para o desenvolvimento do processo metodológico compreende o significado de análise da metodologia Recência, Frequência e Valor (RFV) e a associação da mesma com a leitura dos dados pela abordagem da Lei de Pareto. Entretanto, o arcabouço conceitual que favorece a operacionalização dos dados para a tratativa de análise é exposto pela compreensão dos conceitos fundantes quanto a abordagem da segmentação de clientes para o desenvolvimento de abordagens comerciais específicas que atendam a demanda dos clientes e favoreça satisfação dos mesmos, permitindo à organização ter retenção e fidelização dos clientes (KOTLER, 2000).

O marketing de relacionamento tem como foco o cliente, no intuito de o satisfazer, para garantir uma vivência exitosa que potencialize a fidelização do cliente. Ao estruturar as informações sobre os clientes, a organização amplia sua compreensão sobre a diferença entre os tipos de clientes e os tipos de interações com a organização.

Os clientes são segmentados considerando sua importância e a atividade que exercem nas organizações. É possível pontuar cada cliente e inseri-lo em um determinado intervalo de recência, bem como criar diversos tipos de segmentos baseados na análise RFV (KOTHARI, 2017). Os segmentos identificados no Quadro 1 ilustram algumas possibilidades de desenvolvimento de ações, conforme apontado pelas colunas atividade e dicas.

Tabela 1 - Segmentação de clientes e recomendação de conduta (Kothari 2017).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Segmento de cliente | Pontuação de recência | Intervalo de frequência e valor combinados | Atividade | Ações |
| Campeões | 4-5 | 4-5 | Comprou recentemente, compra sempre e é o que mais gasta. | Recompense-os. Podem ser os primeiros a adotar novos produtos. Irá promover a sua marca. |
| Clientes fiéis | 2-5 | 3-5 | Gasta bastante dinheiro | Venda produtos de maior valor agregado. Peça opiniões. Envolva-os. |
| Potencial cliente fiel | 3-5 | 1-3 | Clientes recentes, mas gastou uma boa quantia e comprou mais de uma vez. | Ofereça programa de fidelidade, recomende outros produtos. |
| Clientes recentes | 4-5 | 0-1 | Compra mais recentemente, mas não com frequência. | Forneça suporte, dê sucesso antecipado e comece a construir um relacionamento. |
| Promissores | 3-4 | 0-1 | Compradores recentes, mas não gastaram muito. | Crie reconhecimento para a marca, ofereça testes gratuitos. |
| Clientes que precisam de atenção | 2-3 | 2-3 | Acima da média de recência, frequência e valores monetários. Porém, podem não ter comprado recentemente. | Faça ofertas com tempo limitado, recomende baseado em compras passadas. Reative-os. |
| Prestes a dormir | 2-3 | 0-2 | Abaixo da média de recência, frequência e valores monetários. Serão perdidos se não forem reativados. | Compartilhe recursos valiosos, recomende produtos populares/renovações com descontos, reconecte-se com eles. |
| Em risco | 0-2 | 2-5 | Gastaram grande dinheiro e compraram com frequência, mas há muito tempo. É preciso trazê-los de volta. | Envie e-mails personalizados para se reconectar, ofereça renovações, providencie recursos úteis. |
| Não pode perdê-los | 0-1 | 4-5 | Fez as maiores compras, com muita frequência. Mas não voltam há tempos. | Ganhe-os de volta com renovações ou produtos recentes, não perca eles pra concorrência, fale com eles. |
| Hibernando | 1-2 | 1-2 | Última compra foi há muito tempo, pouco gasto e poucas vezes que comprou algo. | Ofereça outros produtos relevantes e descontos especiais. Recrie o valor da marca. |
| Perdidos | 0-2 | 0-2 | Menor recência, frequência e valor monetário. | Reative o interesse e ignore caso contrário. |

É importante ressaltar que, de acordo com a Tabela 1, não deve haver a sobreposição de dados, isto é, caso um cliente obtenha a pontuação entre 4 e 5 na recência, mas não ultrapasse 4 na frequência e valor, esse cliente será classificado como um Cliente Fiel. Da mesma forma que, um cliente que possua frequência e valor entre 3 e 5 e sua recência não ultrapasse 4, ele também será classificado como um Cliente Fiel. Logo, o segmento indicado como Campeão é o consumidor que possui ambas as notas entre 4 e 5.

Para Kotler (2000), a concorrência atual não é entre empresas, mas sim entre redes de marketing, destacando-se no mercado a organização que tiver a melhor rede estabelecida. De certa forma, o princípio operacional da concorrência nada mais é do que focar em construir uma rede efetiva de relacionamentos com o público interessado para que os lucros obtidos sejam uma consequência. Com base nessa afirmação, investir no relacionamento com o cliente, bem como no estudo de seu comportamento, se torna essencial em qualquer organização.

De acordo com Sabuncu (2020), RFV pode ser considerado como um método analítico de segmentação comportamental dos clientes. Ainda para esse autor, RFV é uma ferramenta fácil de usar, podendo ser aplicada de forma rápida e é usada continuamente por ser de fácil entendimento. Recência significa a última data de transação do cliente, frequência significa a periodicidade das transações dos clientes e valor se baseia na ideia do dinheiro total que o consumidor gastou. Tal modelo permite descobrir o cliente mais provável de responder a uma nova proposta da empresa, considerando todos os outros consumidores existentes. Clientes que possuem uma alta pontuação de RFV são considerados valiosos para a empresa (SABUNCU, 2020).

# Variáveis utilizadas

Para realização do modelo, foi utilizado os softwares Python para processamento e tratamento do *dataset* primário e inserção no banco de dados, além da modelagem de *machine learning* para validação do resultado. O PostgreSQL foi utilizado para armazenar os dados de forma estruturada, e PowerBI para visualização dos resultados.

O modelo foi criado de forma que os dados são coletados diretamente pelo *script* python, que irá tratar a base de dados de forma automática e efetuar a inserção no banco de dados PostgreSQL. O dado já é atualizado de forma programada no PowerBI, tendo um dashboard sempre atualizado.

# Fontes utilizadas

KOTHARI, Akshat. RFM Analysis For Successful Customer Segmentation. Disponível em: <https://www.putler.com/rfm-analysis/>. Acesso em: 01 de julho de 2023.

KOTLER, Philip. Administração de Marketing: análise, planejamento e controle. 2000. 10ª Edição. Tradução Bazán Tecnología e Lingüística. Revisão técnica Arão Sapiro. São Paulo: Prentice Hall, 2000.

SABUNCU, İbrahim; TÜRKAN, Edanur; POLAT, Hilal. Customer Segmentation and Profiling with RFM Analysis. Turkish journal of marketing, v. 5, n. 1, p. 2236, 2020.

# Dicionário de dados

Para criação do banco de dados foram criadas duas tabelas, rfv.analysis e rfv.evolucao. A tabela rfv.analysis contém as colunas id, rede\_uf, uf, canal, recencia, frequencia, valor, score\_frequencia, score\_recencia, score\_valor e score. A coluna id é a chave primária sequencial (pk), a rede\_uf armazena o nome do cliente, uf refere-se ao estado, canal ao tipo do cliente, recencia data mais recente de compra em mês, frequencia a quantidade de compras do meu cliente, e valor ao valor de compra total dividido pela quantidade de pedidos. Os scores foram os valores atribuídos aos clientes com base nas suas compras. A tabela rfv.evolucao mantem um histórico de todas as colunas de rfv\_analysis com exceção de rede\_uf, uf e canal, que são vinculados através da coluna id\_analysis que é uma chave estrangeira (PFK). Além dessas colunas temos a coluna id que é a chave primária sequencial da coluna rfv.evolucao e data. Nenhum dos campos são aceitos como nulo e a origem dos dados é a base de sell-in da HyperaPharma.

# Cronograma de execução

Tabela 2 – Cronograma de execução

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1° mês | 2° mês | 3° mês | 4° mês | 5° mês |
| Análise dos resultados | X |  |  |  |  |
| Tomada de ações | X | X | X | X | X |
| Inserção periódica dos dados |  | X | X | X | X |
| Análise de evolução |  | X | X | X | X |