

Otimização da Eficiência Logística: Um Estudo de Caso dos Dados de Entrega da Amazon

Introdução

O Dataset de Entregas da Amazon oferece uma visão abrangente das operações de logística de última milha da empresa. Com dados de mais de 43.632 entregas em várias cidades, incluindo informações detalhadas sobre pedidos, agentes de entrega, condições climáticas e de tráfego e métricas de desempenho de entrega, este conjunto de dados permite que pesquisadores e analistas descubram insights sobre os fatores que influenciam a eficiência das entregas, identifiquem áreas para otimização e explorem o impacto de várias variáveis na experiência geral do cliente.

Descrição do Dataset

O dataset é composto por várias colunas, cada uma fornecendo informações cruciais sobre diferentes aspectos das operações de entrega:

- **Agent_Age:** Idade do agente de entrega.
- **Agent_Rating:** Avaliação do agente de entrega.
- **Order_Date:** Data do pedido.
- **Order_Time:** Hora do pedido.
- **Pickup_Time:** Hora da coleta.
- **Weather:** Condições climáticas no momento da entrega.
- **Traffic:** Condições de tráfego no momento da entrega.
- **Area:** Área de entrega.
- **Delivery_Time:** Tempo total de entrega.
- **Category:** Categoria do produto entregue.

Objetivos do Estudo

1. **Identificar fatores que influenciam a eficiência das entregas:** Analisar como variáveis como idade e avaliação do agente, condições climáticas e de tráfego afetam o tempo de entrega.
2. **Descobrir áreas para otimização:** Identificar padrões e gargalos nas operações de entrega que podem ser melhorados.
3. **Explorar o impacto das variáveis na experiência do cliente:** Avaliar como diferentes condições e fatores influenciam a satisfação do cliente.

1 - Perguntar

- Qual é o problema que você está tentando resolver?
 - Otimizar o tempo de entrega das encomendas da Amazon para melhorar a eficiência e a satisfação do cliente.
- Como seus insights podem nortear as decisões de negócios?
 - Insights podem ajudar a identificar os principais fatores que causam atrasos nas entregas, permitindo que a Amazon implemente soluções eficazes para melhorar o desempenho das entregas.

Principais Tarefas

- Identificar a Tarefa de Negócios
 - Analisar os dados de entrega para identificar padrões e gargalos que afetam negativamente o tempo de entrega.
- Considerar as Principais Partes Interessadas
 - Gerentes de logística, motoristas de entrega, equipe de atendimento ao cliente, equipe de análise de dados e clientes da Amazon.

Entrega

- Uma Declaração Clara da Tarefa de Negócios

- "Nosso objetivo é reduzir o tempo médio de entrega em 20% nos próximos seis meses, analisando fatores como condições de tráfego, idade e avaliação dos motoristas, e condições climáticas, a fim de implementar estratégias eficazes de otimização."

2 - Preparar

Os dados estão localizados em um arquivo CSV fornecido pela equipe de logística da Amazon, Os dados são organizados em uma tabela com as seguintes colunas: **Agent_Age**, **Agent_Rating**, **Order_Date**, **Order_Time**, **Pickup_Time**, **Weather**, **Traffic**, **Area**, **Delivery_Time** e **Category**, em que Não há evidência imediata de viés. No entanto, uma análise mais detalhada é necessária para garantir que os dados sejam representativos.

Os dados são originais e abrangentes, cobrindo mais de 43.632 entregas. Foram coletados diretamente do sistema de logística da Amazon, garantindo a atualidade e a fonte confiável. Os dados estão armazenados em um servidor seguro com acesso restrito. Estão em conformidade com as regulamentações de privacidade de dados da Amazon. Realizamos validações para garantir que não há valores ausentes, nulos, duplicados ou inconsistentes nas principais colunas. Aliás, eles fornecem informações detalhadas sobre as entregas, permitindo a análise dos fatores que influenciam o tempo de entrega e a eficiência e que foram identificados alguns outliers nos tempos de entrega e avaliações de agentes que serão tratados na fase de limpeza de dados.

Principais Tarefas

1. **Fazer o Download dos Dados e Armazená-los Adequadamente**
 - i. Os dados foram baixados e armazenados em um banco de dados SQL seguro.
2. **Identificar como Estão Organizados**
 - i. A tabela de dados contém colunas específicas para cada variável relevante: **Agent_Age**, **Agent_Rating**, **Order_Date**, **Order_Time**, **Pickup_Time**, **Weather**, **Traffic**, **Area**, **Delivery_Time** e **Category**.
3. **Classificar e Filtrar os Dados**
 - i. Os dados foram classificados por **Order_Date** e **Delivery_Time**. Registros com dados ausentes ou valores anômalos foram filtrados.
4. **Determinar a Credibilidade dos Dados**
 - i. Os dados foram verificados quanto à origem e método de coleta. Confirmação de que os dados são precisos e confiáveis.

Entrega

1. **Descrição de Todas as Fontes de Dados Usadas**

- i. **Fonte:** Sistema de logística da Amazon
- ii. **Método de Coleta:** Extraído diretamente do site Kaggle
- iii. **Abrangência:** Dados de mais de 43.632 entregas em várias cidades
- iv. **Data de Obtenção:** Atualizado até o mês de junho de 2024

3 - Processar

Estou escolhendo o Power BI para a análise de dados devido à sua capacidade de integrar, limpar, transformar e visualizar dados de maneira eficaz e intuitiva. Além disso, o Power BI permite fácil colaboração e compartilhamento de resultados, tornando-o uma escolha ideal para este projeto. Aliás, a integridade dos dados foi verificada através da remoção de duplicatas, preenchimento de valores ausentes e validação de formatos de dados. Isso assegura que os dados estejam completos e precisos antes de proceder com a análise.

Asseguramos que os dados estejam limpos através da validação de tipos de dados, remoção de valores nulos ou inconsistentes, e normalização dos dados onde necessário. Essas medidas ajudam a garantir a qualidade e a consistência dos dados. A verificação pode ser feita através da inspeção visual dos dados em Power BI, além de checagens automatizadas para valores extremos ou anômalos e revisão de resumos estatísticos. Essas práticas ajudam a identificar e corrigir qualquer problema antes da análise e o processo de limpeza foi documentado em detalhes, incluindo as etapas e transformações realizadas. Essa documentação é essencial para revisão, auditoria e compartilhamento de resultados.

Principais Tarefas

Os dados foram verificados em busca de duplicatas, valores nulos, outliers e inconsistências. Essas verificações são fundamentais para garantir que os dados estejam prontos para análise. Já o Power BI foi escolhido como a ferramenta principal para a importação, limpeza e transformação dos dados, aproveitando suas capacidades de DAX e Power Query. Essa escolha foi baseada na robustez e eficiência do Power BI.

Ademais, os dados foram transformados no Power BI usando o Power Query Editor para remover duplicatas, preencher valores ausentes, normalizar dados e criar colunas calculadas, se necessário. Essas transformações são críticas para preparar os dados para análise. Todo o processo de limpeza foi documentado, incluindo um registro de todas as operações realizadas no Power BI e a documentação é essencial para garantir a transparência e permitir a revisão das etapas de limpeza.

Entrega

Um relatório detalhado foi fornecido, descrevendo todas as operações de limpeza e transformação realizadas nos dados. Este relatório inclui a origem dos dados, as operações realizadas e as justificativas para cada passo, garantindo a clareza e a integridade do processo de limpeza. Tais Operações Realizadas como:

- Remoção de duplicatas nas colunas `Order_ID` e `Delivery_ID`
- Substituição de valores nulos em `Agent_Age` e `Agent_Rating` pela média
- Normalização de formatos de data em `Order_Date` e `Pickup_Time`
- Identificação e tratamento de outliers em `Delivery_Time`

4 - Analisar

Para organizar meus dados de forma eficaz para a análise, agrupei-os por categorias relevantes como área, agente de entrega, condições de tráfego e clima. Os dados foram formatados corretamente, assegurando que todas as variáveis estavam no tipo de dado apropriado e sem valores nulos. Ao explorar os dados, descobri uma correlação entre a idade dos agentes e o tempo de entrega, sugerindo que agentes mais jovens tendem a entregar mais rapidamente. Além disso, notei que entregas em dias nublados tinham um maior tempo de entrega, o que não era inicialmente esperado.

As principais tendências que encontrei foram a relação entre as condições de tráfego e os tempos de entrega, onde tráfego pesado resultava consistentemente em maiores tempos de entrega. Esses insights são cruciais para responder às perguntas de negócios, como identificar áreas para otimização da rota e melhorias na gestão de agentes. Agregando esses dados, podemos tomar decisões informadas para melhorar a eficiência da logística e a satisfação do cliente.

Entrega

Os dados foram organizados e formatados corretamente, revelando surpresas como a correlação entre a idade dos agentes e o tempo de entrega, e a influência negativa das condições climáticas adversas nas entregas. As tendências identificadas, como a relação entre o tráfego pesado e maiores tempos de entrega, são insights valiosos para otimizar a logística e melhorar a eficiência operacional da Amazon.

5 - Compartilhar

Os dados revelam que a eficiência das entregas da Amazon é influenciada por fatores específicos: agentes de entrega mais jovens são mais rápidos, enquanto dependendo das condições de tráfego pesado e clima adverso aumenta ou não o tempo de entrega. Essas descobertas indicam a necessidade de treinamento específico para agentes mais velhos e melhorias no planejamento de rotas para evitar atrasos. Para comunicar esses insights de forma eficaz aos gestores de logística e executivos, utilizei dashboards interativos no Power BI com gráficos de barras para mostrar a performance dos agentes, gráficos de linhas para visualizar tendências de entrega ao longo do tempo, e mapas de calor para ilustrar o impacto do tráfego e clima. Essas visualizações claras e detalhadas garantem uma compreensão completa das áreas críticas que necessitam de otimização, permitindo a tomada de decisões estratégicas fundamentadas.

A análise dos dados revela que a média do tempo de entrega é significativamente maior em áreas semi-rurais. Este insight sugere que a infraestrutura e a acessibilidade nessas regiões podem estar impactando negativamente a eficiência das entregas. Fatores como estradas menos desenvolvidas, maior distância entre os pontos de entrega e menor densidade populacional podem estar contribuindo para esses tempos de entrega mais longos.

A análise revela que a média do tempo de entrega para produtos de mercearia (grocery) é menor. Isso indica que os processos e rotas para entregas de itens perecíveis são mais eficientes, possivelmente devido à prioridade dada a esses produtos para garantir frescor. Este insight sugere que práticas usadas nas entregas de mercearia poderiam ser aplicadas a outras categorias para melhorar a eficiência geral das entregas.

6 - Agir

Com base na análise dos dados de entregas da Amazon, conclui-se que a eficiência das entregas é significativamente influenciada pela idade dos agentes, condições de tráfego e clima, e a localização das áreas de entrega. Descobrimos que agentes mais jovens são mais rápidos, o tráfego pesado e o clima adverso aumentam os tempos de entrega, e que as áreas semi-rurais apresentam maiores desafios logísticos, resultando em tempos de entrega mais longos. Além disso, identificamos que a entrega de produtos de mercearia é mais rápida, possivelmente devido a processos otimizados para itens perecíveis.

Para aplicar esses insights, recomendo que a equipe de logística implemente programas específicos para agentes de entrega mais velhos, focados em técnicas e ferramentas que possam ajudar a melhorar sua eficiência. Além disso, é crucial otimizar as rotas de entrega em áreas com tráfego pesado e clima adverso, utilizando dados em tempo real para planejar as melhores rotas possíveis. Por fim, as práticas eficientes aplicadas nas entregas de mercearia devem ser estudadas e adaptadas para outras categorias de produtos, visando reduzir os tempos de entrega de maneira geral.

Os próximos passos incluem a implementação das recomendações mencionadas, bem como a coleta contínua de dados

para monitorar a eficácia das mudanças realizadas. Seria benéfico explorar dados adicionais sobre a infraestrutura de entrega e a densidade populacional das áreas atendidas para expandir ainda mais as descobertas e identificar outras oportunidades de melhoria. Essas ações ajudarão a Amazon a otimizar suas operações logísticas, melhorar a satisfação do cliente e aumentar a eficiência operacional.

Três Principais Recomendações

1. **Treinamento Específico para Agentes de Entrega Mais Velhos:** Desenvolver e implementar programas de treinamento focados em técnicas e ferramentas que melhorem a eficiência dos agentes de entrega mais velhos, ajudando a reduzir o tempo de entrega e aumentar a produtividade.
2. **Otimização de Rotas em Condições Adversas:** Utilizar dados em tempo real para planejar e otimizar rotas de entrega em áreas com tráfego pesado e condições climáticas adversas, garantindo entregas mais rápidas e eficientes.
3. **Adaptação de Práticas Eficientes de Mercearia para Outras Categorias:** Estudar e adaptar as práticas otimizadas utilizadas nas entregas de produtos de mercearia para outras categorias de produtos, visando reduzir os tempos de entrega e melhorar a eficiência geral da logística.

Essas recomendações, baseadas na análise detalhada dos dados, ajudarão a Amazon a abordar áreas críticas e melhorar significativamente sua operação logística.