

Relatório Detalhado Adventure Works

Nicole Malaquias

1. Introdução

O projeto de modernização de dados da Adventure Works tem como objetivo criar uma infraestrutura robusta e eficiente para atender às demandas analíticas da empresa, com foco em aprimorar a tomada de decisões estratégicas e operacionais. A seguir, serão abordados os tópicos identificados como prioritários durante a consultoria, destacando áreas que podem ser aprofundadas e identificando pontos de atenção que representam potenciais riscos para o sucesso do projeto.

1.1 Objetivos Esperados

- Facilitar o dia a dia da operação e gerência da empresa;
- Apresentar para o CEO, Carlos Silveira, os principais indicadores da Adventure Works, como: Receita Total, Quantidade Total de Pedidos, Produto mais vendido, Cliente que mais comprou, Ranking de todas as lojas;
- Elaborar um Dashboard para o time operacional que seja possível analisar as vendas em 3 recortes: por pedido, por região e por clientes;
- Realizar a previsão da demanda de cada produto em cada loja;
- Descobrir a presença ou não de sazonalidade;
- Definir quais centros de distribuição apresentaram mais crescimento de demanda;
- Estimar a quantidade de zíper a serem solicitados pelo novo fornecedor.

1.2 Partes Interessadas

- Diretoria da Adventure Works;

Carlos Silveira: CEO da empresa;

João Muller: Diretor de Inovação da AW;

Silvana Teixeira: Diretora Comercial;

Nilson Ramos: Diretor de TI;

Luís Soares: Gestor da equipe de Planejamento de Demanda da organização.

- Time de Operações da Adventure Works;
- Novo fornecedor de luvas.

1.3 Riscos e Contingências

Riscos		Contingência
Qualidade dos Dados	Dados podem estar incompletos ou conter valores ausentes	Realizar uma etapa de limpeza e pré-processamento dos dados, estabelecer critérios claros para a remoção de outliers e inconsistências.
Alinhamento com o negócio	O projeto pode não estar alinhado com as metas de negócio, resultando em insights ou soluções que não geram valor.	Incluir stakeholders do negócio no início e em todas as etapas do projeto, validar continuamente se os resultados estão alinhados com as necessidades e objetivos da empresa.
Custo e Tempo para Implementação	O tempo necessário para a realização de todo o processo não pode ser subestimado, corre-se o risco de ultrapassar o orçamento	Estabelecer cronogramas com marcos bem definidos. É necessário um monitoramento regular sobre o progresso e os custos para que grandes desvios sejam evitados.
Falta de	Os stakeholders podem não entender	É fundamental estabelecer uma

compreensão dos benefícios do projeto	o valor que o projeto agrega ao negócio. Podendo gerar atrasos e comprometer a qualidade dos resultados.	comunicação clara com todos os stakeholders para alinhar as expectativas. Definir claramente os objetivos, metas e o valor que a análise de dados e a ciência de dados podem agregar à empresa. Além disso, é importante definir KPIs (indicadores-chave de desempenho) que permitam medir o sucesso e o impacto do projeto ao longo do tempo.
Prioridades conflitantes	A empresa contratada pode ter outras prioridades estratégicas ou operacionais que possam ofuscar a implementação do projeto. Podendo interromper ou atrasar o progresso do projeto.	Manter um canal de comunicação aberto e contínuo com os stakeholders. Realizar reuniões de acompanhamento regulares para discutir os progressos, os desafios e as próximas etapas do projeto, criando um ambiente de transparência.

2. Estratégia de Implementação

Para implementar o projeto e criar uma infraestrutura de dados eficiente, fundamental para o sucesso de qualquer iniciativa data-driven, optou-se pelo Modern Data Stack (MDS). Esta abordagem se baseia em boas práticas e ferramentas modernas, criando uma arquitetura de dados flexível, escalável, autônoma e alinhada com as tendências atuais.

O MDS facilita a integração de diversas ferramentas, cada uma com funções específicas, resultando em uma infraestrutura moderna. Durante o planejamento da implementação foi idealizado fases de execução e responsabilidades claras, garantindo uma distribuição eficaz das tarefas e a agilidade necessária para que as equipes de dados funcionem como times de produto.

Essa abordagem não só evita o lock-in e assegura a portabilidade dos dados, como também democratiza o acesso aos dados em toda a organização. A arquitetura MDS é projetada para ser cloud-native, proporcionando escalabilidade, flexibilidade e redução de custos com infraestrutura e manutenção. Por essas razões, o MDS foi escolhido como a base da transformação data-driven da **Adventure Works**, posicionando a empresa para responder eficazmente às crescentes demandas analíticas e operacionais.

2.1 Aquisição, Integração e Tratamento de Dados

Durante o diagnóstico inicial, foram identificados sistemas essenciais que geram dados relevantes para o negócio, como o SAP, utilizado para o gerenciamento do ERP, o Salesforce, que apoia o CRM, o Google Analytics e o Wordpress. Devido ao prazo para a execução do projeto, a opção mais ágil foi utilizar arquivos CSV

disponíveis desses sistemas. Esses CSVs foram carregados no data warehouse utilizando seeds do dbt, evitando a etapa de ingestão tradicional, que demandaria mais tempo.

2.2 Estruturação e Gestão de Dados

Para o armazenamento e gerenciamento dos dados, foi utilizado o BigQuery, uma plataforma de data warehouse altamente escalável e eficiente da Google Cloud. A escolha pelo BigQuery se deu pela sua capacidade de processar grandes volumes de dados com rapidez e pela facilidade de integração com ferramentas como o dbt, permitindo consultas SQL otimizadas e a execução de análises em tempo real. Além disso, o BigQuery oferece flexibilidade e escalabilidade, essenciais para acompanhar o crescimento acelerado da Adventure Works, garantindo que o ambiente de dados possa evoluir junto com as necessidades do negócio.

2.3.1 Processamento e Refinamento de Dados

Para otimizar o tempo de desenvolvimento do projeto, foi adotada a abordagem ELT (Extract-Load-Transform), na qual os dados foram extraídos, carregados e posteriormente transformados. A ferramenta principal utilizada para a transformação dos dados foi o dbt. Em vez de realizar a ingestão tradicional dos dados, optou-se por utilizar o recurso de *dbt seeds*, permitindo o carregamento direto de arquivos CSV no data warehouse.

2.3.2 Preparação dos Dados para Consumo em BI e Definição do

Dataset

A preparação dos dados para consumo em Business Intelligence (BI) e a definição do dataset foram realizadas em três etapas principais: staging, intermediate e marts. Cada uma dessas camadas tem um papel específico no processo de transformação e organização dos dados, garantindo que eles estejam prontos para análise e geração de insights estratégicos.

Na etapa de staging, a principal preocupação foi garantir a seleção, limpeza e formatação inicial dos dados. Foram realizados processos como a seleção e renomeação de colunas, conversão de tipos de dados (cast), achatamento de objetos estruturados e aplicação de filtros iniciais. Além disso, foram executados testes para validar a integridade das fontes de dados. Durante essa fase, foi definida a criação de queries ephemeral para evitar o armazenamento desnecessário e reduzir custos operacionais, especialmente em ambientes como o BigQuery, onde o processamento de um volume de dados muito grande pode não ser interessante financeiramente para um começo de projeto.

Em seguida, na camada intermediate, os dados foram preparados para direcionamento aos marts. Essa etapa envolveu a combinação de diferentes tabelas de staging, aplicação de funções complexas, agregações e a criação de métricas e regras de negócio.

Por fim, na etapa de marts, os dados foram organizados em um modelo dimensional, seguindo o esquema estrela (star schema).

Essa estruturação em camadas e o cumprimento das boas práticas de codificação e testes garantiram que os dados fossem devidamente preparados para consumo em BI, oferecendo uma base sólida para análises detalhadas e precisas.

2.4 Análise Estratégica

Para possibilitar a visualização dos dados, realizar análises detalhadas e gerar insights valiosos, foi utilizada a ferramenta Power BI. A escolha pelo Power BI se

deu pela sua capacidade de integrar-se facilmente com o data warehouse e transformar grandes volumes de dados em dashboards interativos e intuitivos.

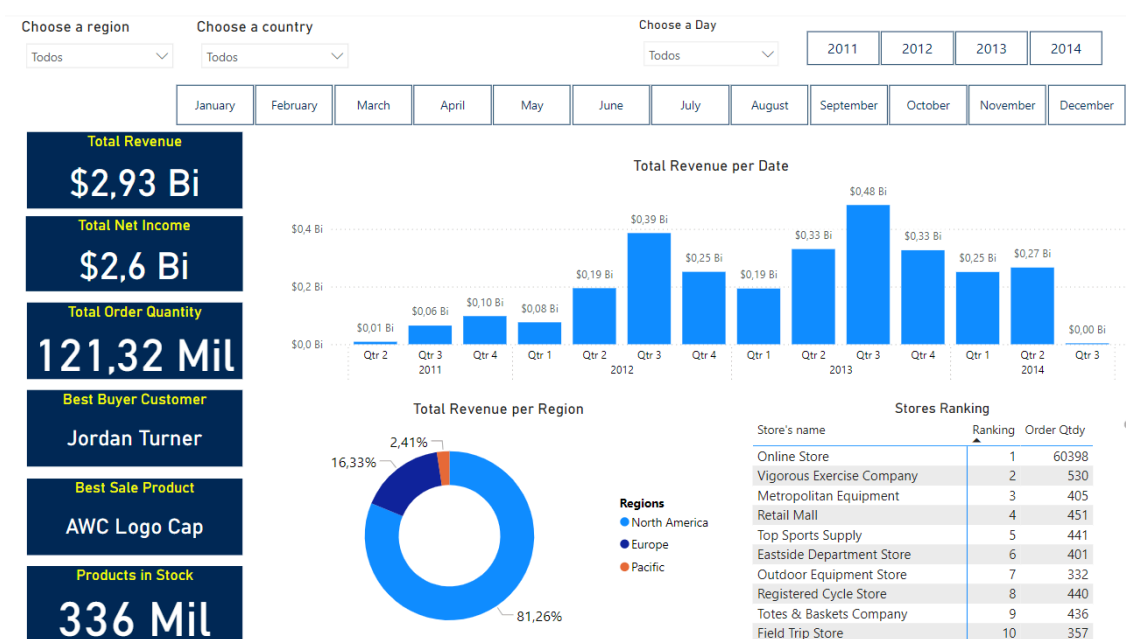
2.5 Análises Avançadas e Modelagem Preditiva

Para realizar a manipulação dos dados no processo de predição de demanda, utilizei a biblioteca Pandas. Com suas estruturas de dados versáteis, como DataFrames, o Pandas permitiu organizar, limpar e transformar os dados de forma rápida e eficiente, preparando-os para as etapas subsequentes de modelagem e previsão. Além disso, para lidar com as séries temporais hierárquicas, utilizei a ferramenta Hierarchical Séries da Nixtla, que foi executada dentro do ambiente Google Colab.

3. Análise dos Dados

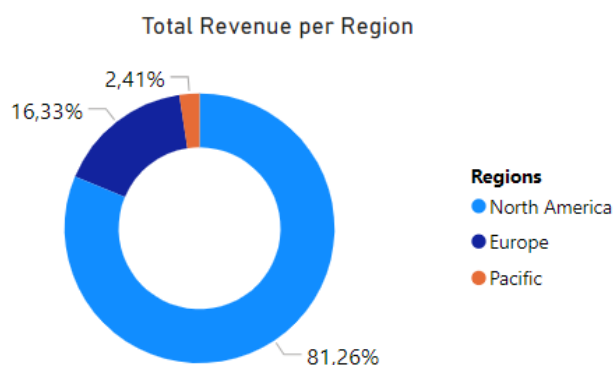
3.1 Relatório Gerencial

Foi elaborado o Relatório Gerencial cujo objetivo é apresentar os principais indicadores a serem vistos diariamente pelo CEO, Carlos Silveira. Foi utilizado informações agregadas, a fim de gerar insights rápidos para a gerência da empresa.



Na região esquerda do relatório, acompanhando o modo como lê-se no ocidente, têm-se os indicadores: Total Revenue (Receita Total), Total Net Income (Receita Líquida Total), Total Order Quantity (Quantidade Total de Pedidos), Best Buyer Customer (Cliente que mais comprou), Best Sale Product (Produto mais vendido), Products in Stock (Quantidade de produtos em estoque).

No Gráfico de Rosca Total Revenue per Region (Receita Total por Região) é notório que a região com mais vendas é a América com Norte com 81% de todas as vendas. Seguida pela Europa com 16% e por último a Oceania com apenas 2% do Total das Receitas.

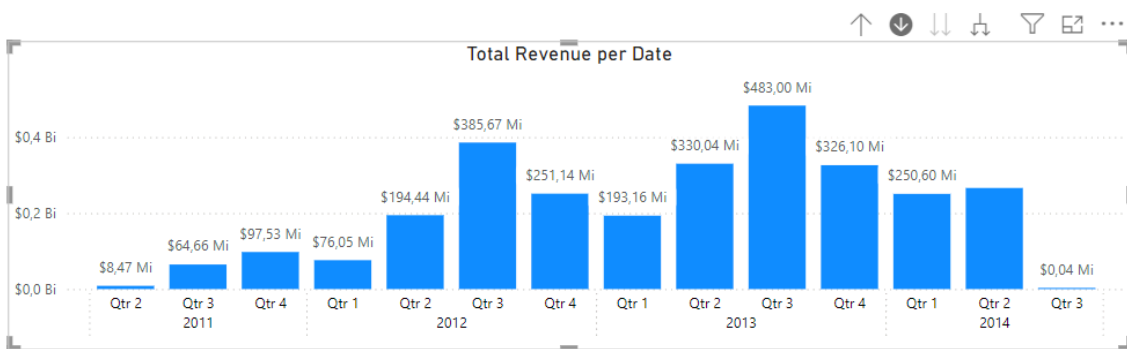


Na Tabela de Stores Ranking (Ranking de Lojas), têm-se rapidamente o Top 10 das lojas que mais venderam em todos os tempos. Caso os filtros sejam aplicados é possível notar a variação desse ranking. Quando o Order Qty (Quantidade de Pedidos) estiver como 0, significa que não houve pedidos para aquela loja no período selecionado.

Stores Ranking

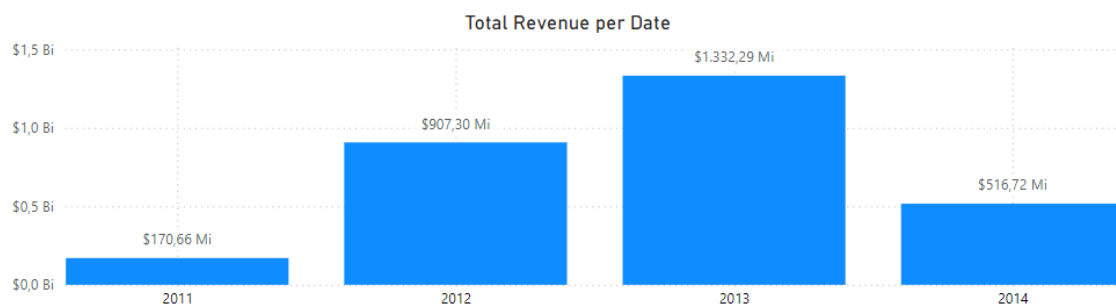
Store's name	Ranking	Order Qty
Online Store	1	60398
Vigorous Exercise Company	2	530
Metropolitan Equipment	3	405
Retail Mall	4	451
Top Sports Supply	5	441
Eastside Department Store	6	401
Outdoor Equipment Store	7	332
Registered Cycle Store	8	440
Totes & Baskets Company	9	436
Field Trip Store	10	357

No Gráfico de colunas do Total Revenue per Date (Receita Total por Data) é possível analisar periodicamente o valor da Receita para as datas. Observa-se o crescimento das vendas em 2012 e 2013.

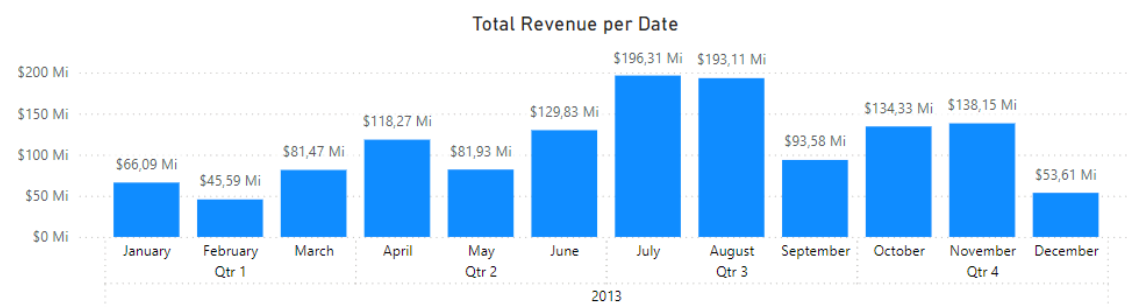


Para uma análise mais detalhada do período, é possível utilizar o Drill up and Down para variar entre a visão geral (acima) e entre Ano, Quarter, Mês e Dia.

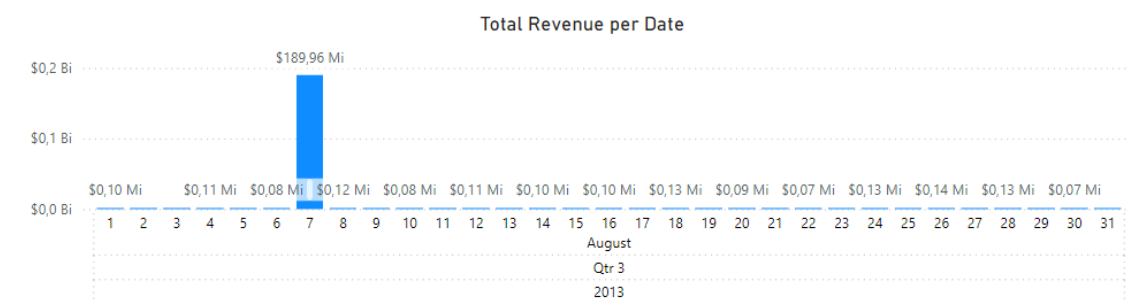
Por anos:



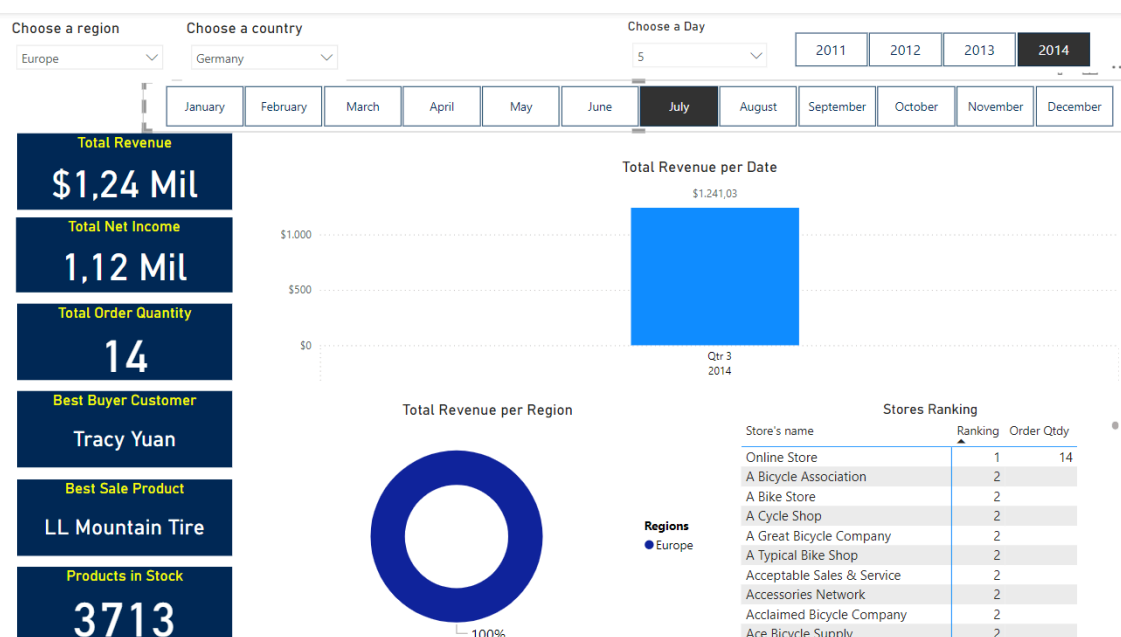
Clicando duas vezes na coluna 2013:



Clicando duas vezes em Agosto:



É possível utilizar os filtros de Região, País, Dia, Mês e Ano para ter uma visão detalhada:

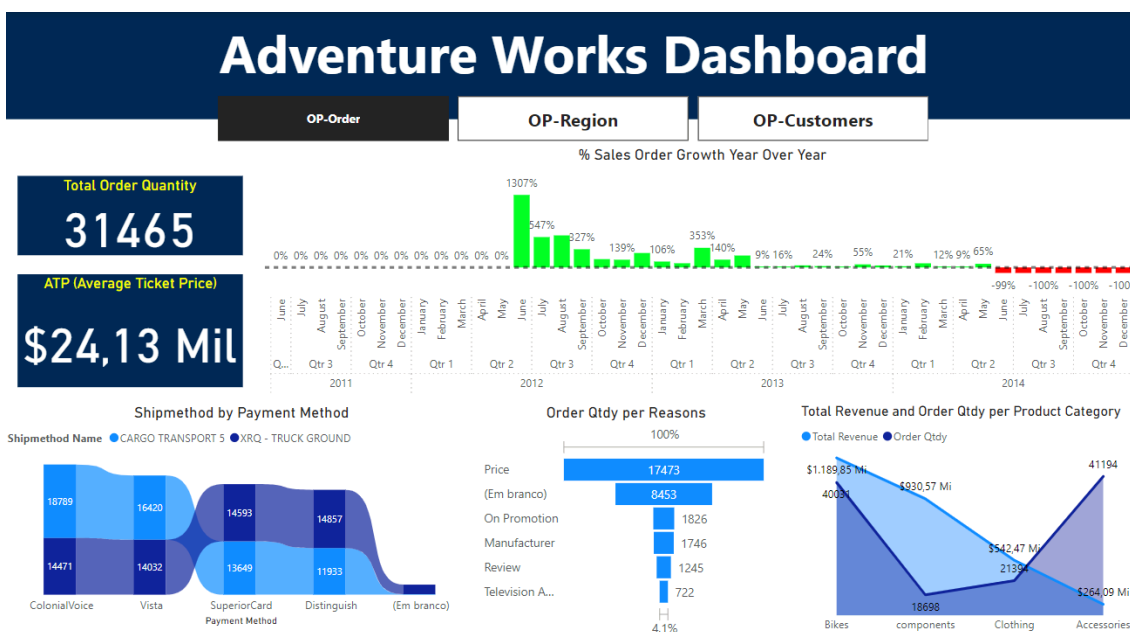


Nessa visão observa-se que a Região escolhida foi Europe (Europa), o país foi Germany (Alemanha), no dia 05 de Julho de 2014. Nota-se que apenas houveram vendas Online, foram 14 Pedidos feitos, que geraram uma receita de \$ 1241,03 e uma receita líquida de \$ 1,12 Mil. O Cliente que mais comprou foi Tracy Yuan, o produto mais vendido foi LL Mountain Tire. E nesse período existiam 3713 produtos em estoque.

3.2 Dashboard de Operações da AW

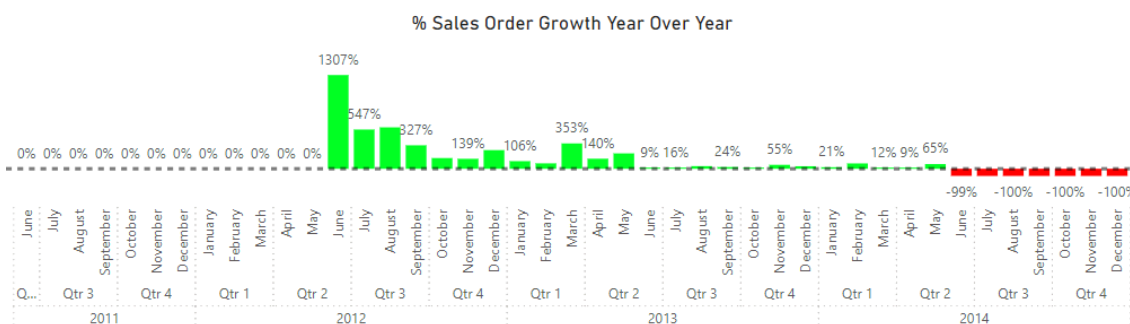
O Dashboard de Operações é dividido em três visualizações: OP-Order, OP-Region, OP-Customers. OP sendo a abreviação de OPerações. Todas as opções nos eixos apresentadas são clicáveis, o que torna possível a visualização personalizada com filtros diversos.

Na visualização de OP-Order, tem-se uma apresentação com maior granularidade sobre os Pedidos.



Na região logo abaixo de “Adventure Works Dashboard” possui botões clicáveis que redirecionam a visualização para o Dash correspondente.

Na visualização de Pedidos tem-se dois indicadores em cartões a esquerda, Total Order Quantity (Quantidade Total de Pedidos) e ATP - Average Ticket Price (Ticket Médio). Em seguida na região central e ocupando a maior parte da visualização tem-se o % Sales Order Growth Year Over Year (Porcentagem de Crescimento dos Pedidos de Vendas Year Over Year).

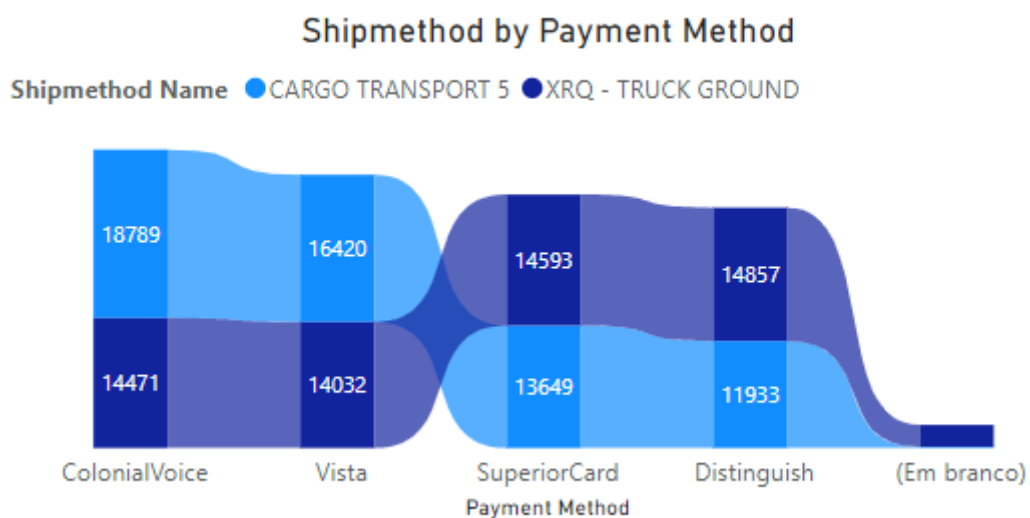


Year over year (YoY) é um cálculo utilizado para analisar a saúde da companhia. Ela permite avaliar se o desempenho financeiro de uma empresa está melhorando, piorando ou se está estagnado. No gráfico de barras 2011 está zerado, pois não existem informações de 2010 para comparar, 2011 é o primeiro ano de nossa

análise. Já as informações após Agosto de 2014 podem ser descartadas já que o banco de dados possui informações até Julho de 2014.

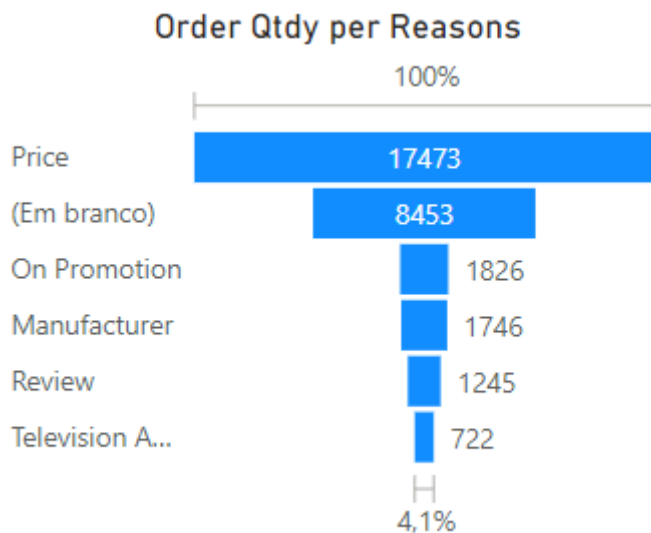
Em 2012, principalmente, observa-se um grande crescimento percentual da Adventure Works, em 2013 e 2014 esse crescimento está estável. Porém em Junho e Julho de 2014, nota-se que o crescimento foi negativo, é necessário investigações mais profundas para entender os motivos dessa desaceleração no crescimento. Seria ideal observar se os próximos meses de 2014 manterão essa desaceleração ou se esse foi um caso isolado.

No Gráfico de Faixas abaixo, tem-se a distribuição pelo método de pagamento.



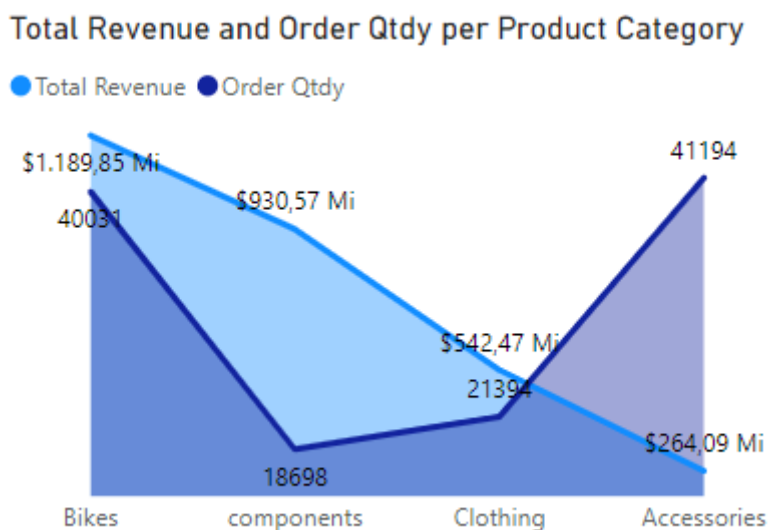
Nela fica evidente que houve apenas 2 meios de distribuição: Cargo Transport 5 e XRQ - Truck Ground. O Cargo Transport 5 foi mais utilizado em vendas com o método de pagamento ColonialVoice e Vista. Enquanto que o método XRQ foi mais utilizado para os pagamentos com SuperiorCard e Distinguish. Em alguns casos o método de pagamento não foi informado, por isso a presença do campo Em Branco no gráfico.

No Gráfico de Funil abaixo, tem-se a Quantidade de Pedidos pelas razões.



Nesse gráfico é ilustrado que o principal motivo de compra foi o Preço, seguido por um grande percentual em que não foi mencionado o motivo. Em terceiro temos Produtos em Promoção. Com isso fica evidente que o que mais atrai os consumidores da Adventure Works é o Preço.

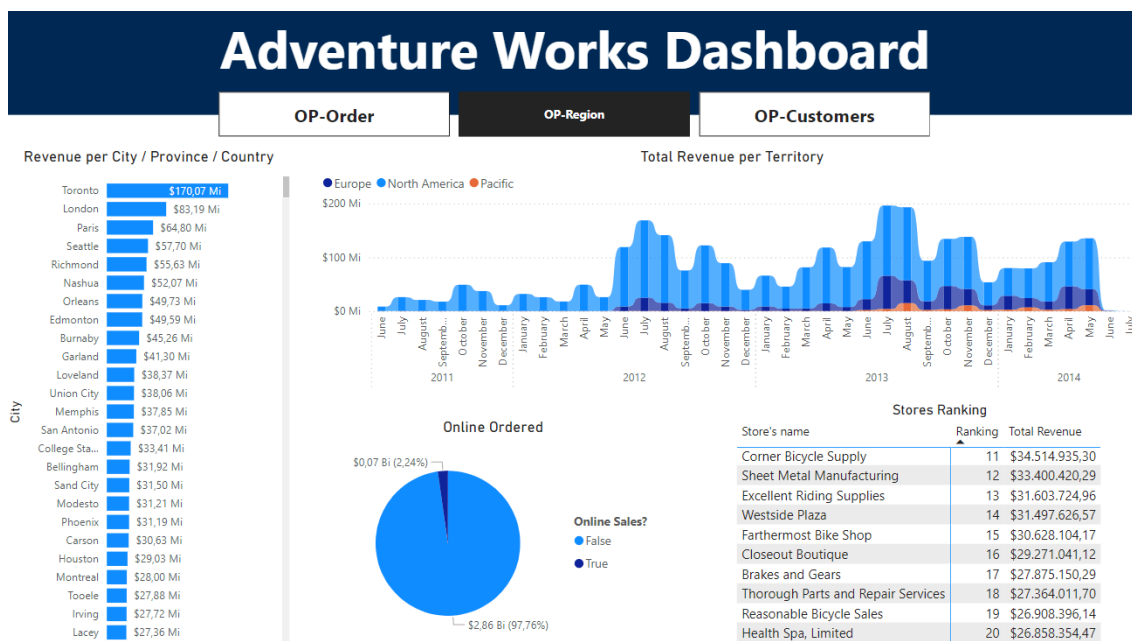
No Gráfico de área ilustrado abaixo, tem-se a Receita Total e a quantidade de pedidos pelas Categorias dos Produtos.



Analisando o gráfico, tem-se que os Acessórios foram os Produtos mais vendidos em relação a quantidade, porém a Receita gerada por eles foi inferior a todas as

outras. Já as bicicletas foram os produtos com a maior Receita e a segunda maior quantidade. As bicicletas e os Componentes são as duas principais categorias vendidas pela Adventure Works.

Na visualização de OP-Region, tem-se uma apresentação com maior granularidade sobre a Região.

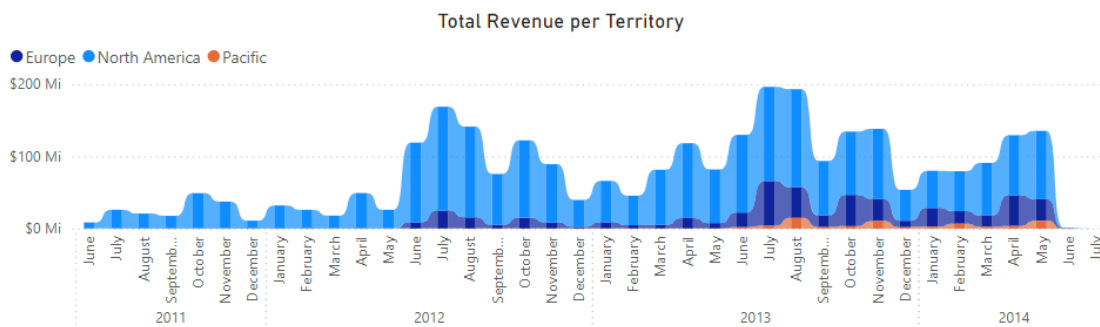


A seguir será explorado cada um dos itens mencionados na Visualização. É possível clicar em itens dos eixos individualmente e observar como todo o Dashboard se ajusta a esse novo filtro.



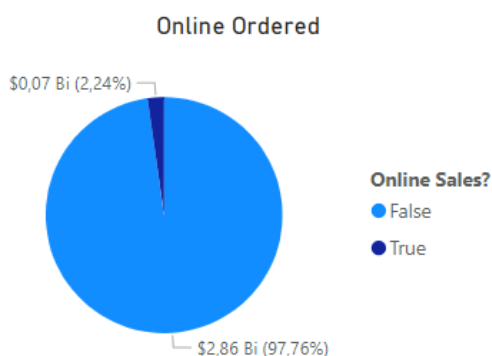
No Gráfico de Barras ao lado, tem-se a Receita por Cidade, Província e País. As localizações podem ser alternadas utilizando os botões de Drill Up e Down. Com essa possibilidade tem-se rapidamente quais são as localidades com as maiores Receitas.

O Gráfico de Faixas abaixo ilustra a Receita Total por Território.



Essa apresentação é apresentada no decorrer dos meses. Em 2011 a Adventure Works só possuía mercado na América do Norte, em 2012 expandiu o seu mercado para a Europa e em 2013 iniciou-se as vendas na Oceania. É notório que a América do Norte é o local com a maior concentração das vendas da AW.

O Gráfico de Pizza ao lado demonstra os pedidos que foram feitos Online ou Não.



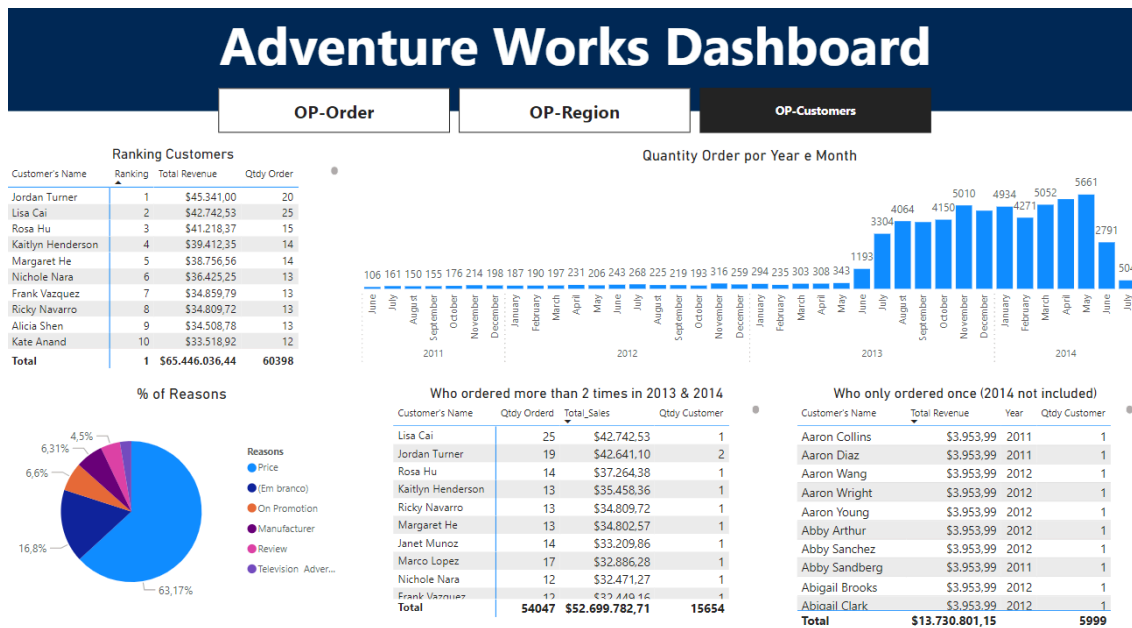
Nele percebe-se que 97% das vendas são realizadas em lojas físicas. Enquanto apenas 2% são feitas em lojas online. Existe uma chance de explorar mais o mercado online para tentar atrair outros tipos de consumidores.

Por fim, no Dashboard de OP-Region tem-se a Tabela de Stores Ranking (Ranking das Lojas).

Stores Ranking		
Store's name	Ranking	Total Revenue
Online Store	1	\$65.446.036,44
Vigorous Exercise Company	2	\$41.901.163,23
Metropolitan Equipment	3	\$38.882.595,01
Retail Mall	4	\$38.595.035,72
Top Sports Supply	5	\$38.426.845,59
Eastside Department Store	6	\$38.056.813,11
Outdoor Equipment Store	7	\$38.019.405,04
Registered Cycle Store	8	\$37.509.705,56
Totes & Baskets Company	9	\$36.244.446,96
Field Trip Store	10	\$35.967.231,44

Nela é possível observar o ranking das lojas juntamente com a Receita Total de cada uma delas. Algumas dessas lojas estão presentes em mais de uma região, como a Loja Online, que é a número 1 em vendas.

Na visualização de OP-Customers, tem-se uma apresentação com maior granularidade sobre os Consumidores.



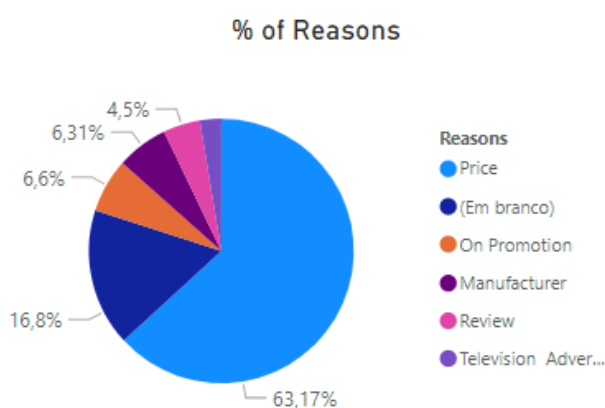
A seguir será explorado cada um dos itens mencionados na Visualização. É possível clicar nos nomes dos Clientes individualmente e verificar como todo o Dashboard se altera para corresponder ao que foi selecionado como filtro. É possível clicar em qualquer opção listada nos eixos ou legendas para ter uma visualização específica. Iniciando a análise de todos os itens pelo lado superior esquerdo, tem-se a Tabela Ranking dos Consumidores.

Customer's Name	Ranking	Total Revenue	Qty Order
Jordan Turner	1	\$45.341,00	20
Lisa Cai	2	\$42.742,53	25
Rosa Hu	3	\$41.218,37	15
Kaitlyn Henderson	4	\$39.412,35	14
Margaret He	5	\$38.756,56	14
Nichole Nara	6	\$36.425,25	13
Frank Vazquez	7	\$34.859,79	13
Ricky Navarro	8	\$34.809,72	13
Alicia Shen	9	\$34.508,78	13
Kate Anand	10	\$33.518,92	12
Total	1	\$65.446.036,44	60398

Customer's Name	Ranking	Total Revenue	Qty Order
Abigail Bennett	18390	\$2,53	1
Abigail Morris	18390	\$2,53	1
Alex Collins	18390	\$2,53	1
Brad Kumar	18390	\$2,53	1
Cody Sanders	18390	\$2,53	1
Dylan Taylor	18390	\$2,53	1
Hunter Miller	18390	\$2,53	1
Marcus Morgan	18390	\$2,53	1
Melanie Peterson	18390	\$2,53	1
Natalie Bryant	18390	\$2,53	1
Total	1	\$65.446.036,44	60398

Na visualização inicial da Tabela tem-se os 10 principais consumidores. Esse ranking foi criado com base na Receita Total gerada pelas compras. A quantidade de pedidos ajuda a perceber com que frequência esses consumidores realizam novas compras na AW. É possível clicar em cima de "Ranking" e ordenar inversamente essa tabela, com isso será apresentado as informações dos consumidores que menos gastaram nas lojas da AW, como ilustrado na figura da direita.

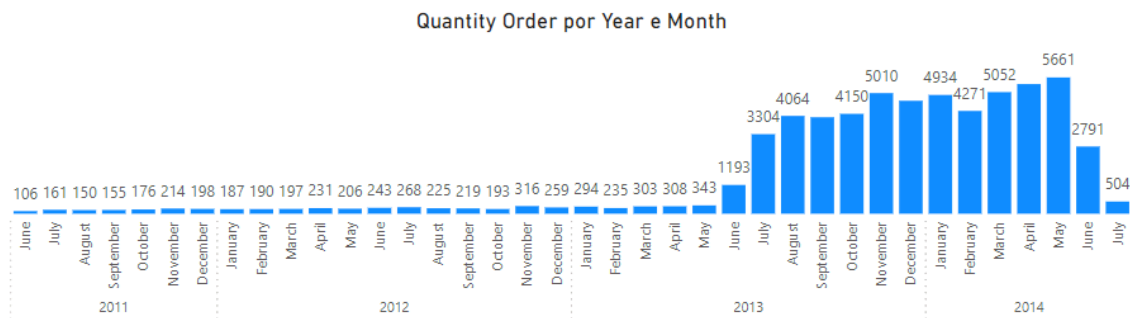
A seguir tem-se o Gráfico de Pizza com a Porcentagem de Razões por pedido de compra.



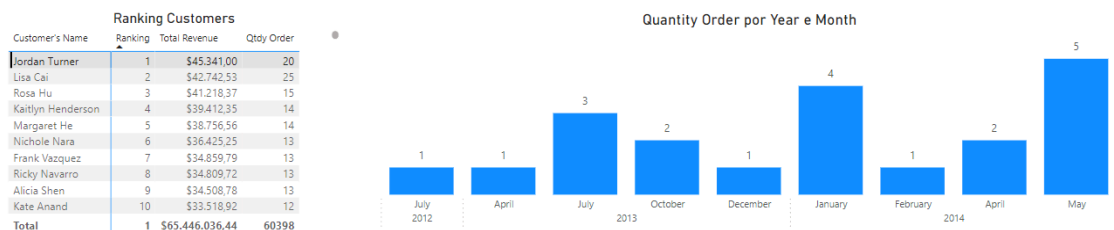
Nota-se que 63% das compras realizadas teve como principal motivo de compra o Valor. Não tiveram um motivo informado, está apresentado como "Em Branco". Seguido, em terceiro, por produto em promoção. O que mais atrai os

consumidores para as lojas da Adventure Works são os preços diferenciados.

O Gráfico de Quantidade de Pedidos dos Clientes por Ano e Mês está ilustrado abaixo.



É notório que a quantidade de pedidos cresceu a partir do segundo semestre de 2013 até Maio de 2014. O principal motivo desse gráfico estar presente nessa página, é para possibilitar a rápida visualização de períodos em que o cliente realizou compras na loja.



No caso acima o nome do Jordan Turner está selecionado, e observa-se que ele realizou mais compras em 2013 e 2014.

A Tabela de quem comprou mais do que duas vezes em 2013 e 2014, tem como principal intuito verificar os clientes que mais compraram nos últimos 2 anos.

Who ordered more than 2 times in 2013 & 2014

Customer's Name	Qty Orderd	Total Sales	Qty Customer
Lisa Cai	25	\$42.742,53	1
Jordan Turner	19	\$42.641,10	2
Rosa Hu	14	\$37.264,38	1
Kaitlyn Henderson	13	\$35.458,36	1
Ricky Navarro	13	\$34.809,72	1
Margaret He	13	\$34.802,57	1
Janet Munoz	14	\$33.209,86	1
Marco Lopez	17	\$32.886,28	1
Nichole Nara	12	\$32.471,27	1
Frank Vazquez	12	\$32.440,16	1
Total	54047	\$52.699.782,71	15654

Com essa informação disponível, é possível direcionar e-mails de marketing para esse público em específico. A última coluna da tabela é para gerar um cálculo rápido de quantos Consumidores estão sendo listados.

E por fim, a Tabela de Quem realizou apenas uma compra em todos esses anos de atividade da Adventure Works, com exceção do ano de 2014.

Who only ordered once (2014 not included)			
Customer's Name	Total Revenue	Year	Qty Customer
Aaron Collins	\$3.953,99	2011	1
Aaron Diaz	\$3.953,99	2011	1
Aaron Wang	\$3.953,99	2012	1
Aaron Wright	\$3.953,99	2012	1
Aaron Young	\$3.953,99	2012	1
Abby Arthur	\$3.953,99	2012	1
Abby Sanchez	\$3.953,99	2012	1
Abby Sandberg	\$3.953,99	2011	1
Abigail Brooks	\$3.953,99	2012	1
Abigail Clark	\$3.953,99	2012	1
Total	\$13.730.801,15		5999

O foco dessa tabela é direcionar e-mails de marketing para as pessoas que estão sem comprar há algum tempo e que tiveram apenas uma compra feita.

É possível filtrar por ano e direcionar o foco do e-mail de acordo com quantos anos fazem desde a última compra.

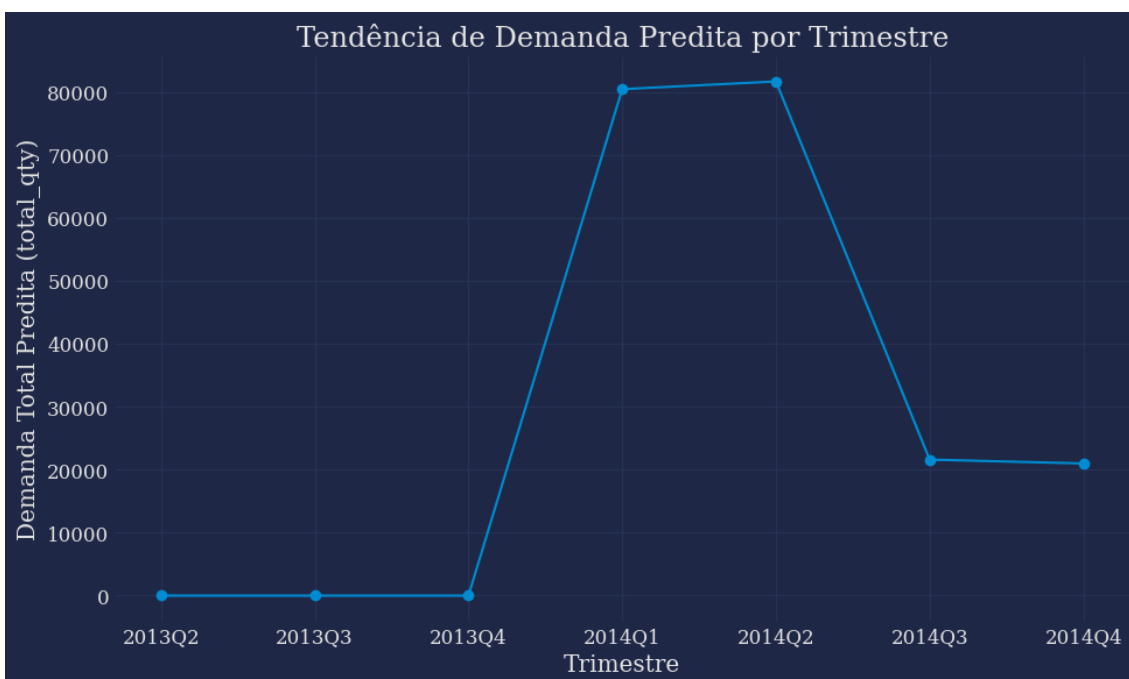
Como na Tabela anterior, essa também contém na última coluna a quantidade de Consumidores presentes nesta listagem.

4. Previsão Trimestral de Demanda e Análise de Sazonalidade

Para o planejamento da demanda da organização, foi realizada uma análise utilizando o modelo Auto ARIMA para determinar os melhores parâmetros do modelo ARIMA. Com base nessa análise, a previsão da demanda foi feita para os próximos três meses, considerando cada loja e produto vendido pela Adventure Works no período disponível.

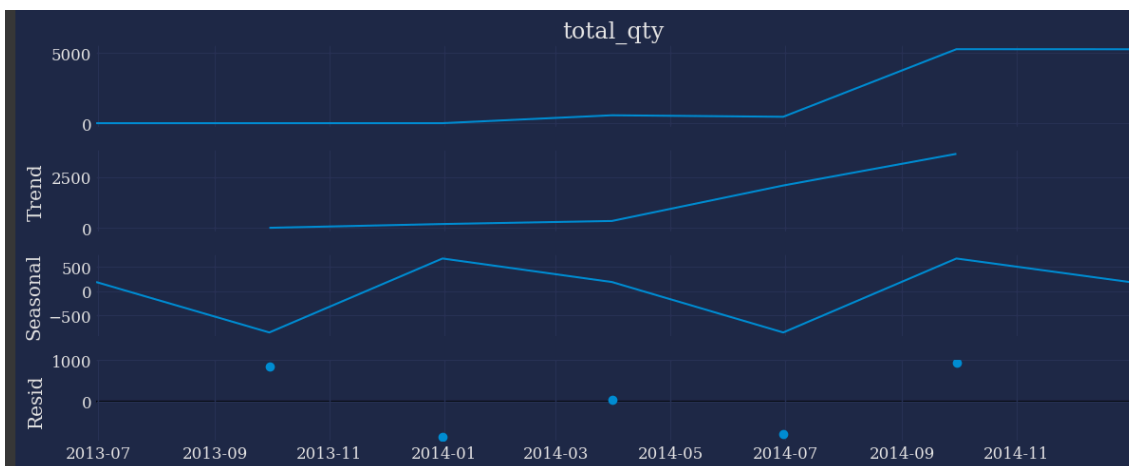
Devido a limitações de poder computacional, foram necessários alguns ajustes para viabilizar a previsão para um grande número de lojas e produtos. Apenas as vendas de lojas cadastradas na base de dados foram consideradas, e durante o período de análise, observou-se que muitos produtos, embora diferenciados por detalhes menores, compartilhavam o mesmo número de produto e custo. Esses produtos foram agrupados para otimizar o processo de previsão, garantindo uma análise eficiente sem comprometer a precisão dos resultados.

Apesar dessas otimizações, o treinamento do modelo ainda exigia algumas horas, e, em alguns momentos, o ambiente do Colab não conseguia lidar com o volume de dados. Como solução, optou-se por agrupar as demandas trimestralmente, reduzindo os 6 milhões de registros originais para cerca de um quarto desse número. Essa abordagem tornou a previsão mais viável dentro das limitações computacionais, permitindo que o processo fosse concluído de maneira mais eficiente.



A imagem acima apresenta a previsão para o quarto trimestre de 2014, onde é possível observar que a previsão sugere uma possível estabilidade nas vendas gerais. Isso indica que, após um período de crescimento, as vendas podem se manter constantes, sem grandes variações, ao longo desse período.

Também foi elaborado o gráfico abaixo para realizar uma análise da sazonalidade nas vendas do produto com maior fluxo de demanda, o Touring-1000. A análise revelou que o produto apresenta padrões sazonais bem definidos, evidenciando variações previsíveis em determinados períodos do ano.



Para realizar a análise de sazonalidade, foram utilizados tanto os dados históricos da base de dados quanto as previsões geradas pelo modelo de machine learning.

Métricas

Level	Metric	ARIMA/BottomUp
Overall	mse-scaled	137.423
	mae-scaled	635.186
	rmse-scaled	370.707
country_name	mse-scaled	133.733
	mae-scaled	388.331
	rmse-scaled	365.695
country_name/province_name	mse-scaled	138.619
	mae-scaled	456.019
	rmse-scaled	372.316
country_name/province_name/store_name	mse-scaled	1.313.452
	mae-scaled	1.826.008
	rmse-scaled	1.146.059
country_name/province_name/store_name/product_name	mse-scaled	1.0
	mae-scaled	1.0
	rmse-scaled	1.0

Valores baixos nessas métricas indicam que o modelo conseguiu prever a demanda com precisão, refletindo bem as variações sazonais ao longo do tempo. Em outras palavras, significa que o modelo foi eficaz em entender os padrões de aumento e queda na demanda, permitindo um planejamento mais acertado.

Por outro lado, valores altos nessas métricas sugerem que o modelo enfrentou dificuldades, resultando em previsões menos precisas e possivelmente falhas em capturar corretamente a sazonalidade.

Um fator que afetou negativamente as previsões foi a falta de dados completos sobre a demanda de todos os produtos em todas as lojas. A ausência dessas informações limitou a capacidade do modelo de capturar plenamente os padrões

sazonais, o que comprometeu a acurácia das previsões. Para melhorar a qualidade das previsões futuras, é crucial garantir que a base de dados esteja completa e contenha informações detalhadas sobre todos os produtos.

Em resumo, as métricas mostram que o modelo ARIMA pode ser eficaz na previsão de demanda e na captura da sazonalidade, mas sua performance pode ser prejudicada pela falta de dados completos, evidenciando a necessidade de uma base de dados robusta para previsões mais confiáveis.

5. Análise de Crescimento de Demanda por Região: Províncias dos EUA vs. Países Internacionais

A análise dos dados de demanda para os três meses previstos indicou que a província do Colorado, nos Estados Unidos, foi a região que apresentou o maior crescimento. O Colorado registrou um aumento excepcional, com uma mudança percentual de 112.282.107,56%, destacando-se significativamente entre as demais regiões analisadas.

Em comparação, a França, que também apresentou um crescimento notável, alcançou uma mudança percentual de 47.213.183,23%. Apesar desse aumento expressivo, ele ainda ficou aquém do crescimento observado no Colorado.

Para calcular esses percentuais de crescimento, utilizamos como base os dados de demanda dos três meses anteriores aos três meses previstos. A diferença percentual foi então calculada comparando-se a demanda predita com a demanda observada nesse período anterior.

Com base nesses resultados, é evidente que a província do Colorado se destacou como a região com o maior crescimento de demanda, superando não apenas outras províncias dos EUA, mas também os países fora dos EUA analisados no período.

6. Avaliação da Solução via Modelos de Regressão e Comparação de Desempenho

Nesta análise, comparamos dois modelos – ARIMA HTS e um modelo de regressão – para prever a demanda dos próximos três meses. Para garantir uma comparação justa, utilizamos os mesmos dados para ambos os modelos, permitindo avaliar qual deles é mais adaptável aos dados disponíveis.

Modelo ARIMA HTS:

- **MSE (Mean Squared Error):** O ARIMA HTS apresentou um MSE geral de 1340.501429. Valores menores de MSE indicam melhor desempenho do modelo, sugerindo que o ARIMA HTS consegue capturar de maneira eficiente as variações na demanda, mesmo nos níveis mais detalhados da hierarquia.
- **MAE (Mean Absolute Error):** O MAE no nível mais detalhado (produto em cada loja) foi de 9496.168781. Embora este valor seja relativamente alto, ele é compreensível dado o nível de granularidade e a complexidade envolvida em prever a demanda a esse nível.
- **RMSE (Root Mean Squared Error):** O RMSE geral foi de 36.612859, o que indica que o modelo ARIMA HTS tende a errar, em média, por cerca de 37 unidades. Este erro é aceitável, especialmente considerando a capacidade do modelo em lidar com séries temporais complexas e hierárquicas.

Modelo de Regressão:

- **R² Score:** O modelo de regressão apresentou um R² negativo (-0.0050), o que indica que ele não foi eficaz em explicar a variabilidade dos dados. Um R² negativo sugere que o modelo de regressão pode ser menos eficaz do que uma simples previsão baseada na média.

- **MAE:** O MAE do modelo de regressão foi de 0.693494685897352, o que parece baixo comparado ao ARIMA HTS. No entanto, esse valor pode ser enganoso devido ao nível de agregação dos dados. A agregação pode ocultar a complexidade e suavizar variações que o ARIMA HTS é capaz de capturar.
- **MAPE (Mean Absolute Percentage Error):** O MAPE extremamente alto (547649709567662.1) indica que o modelo de regressão está cometendo grandes erros percentuais, especialmente em casos onde a demanda é baixa. Isso sugere que o modelo de regressão pode estar subestimando ou superestimando drasticamente a demanda em situações específicas.

Embora seja possível utilizar um modelo de regressão para resolver o problema de previsão de demanda, os resultados indicam que o ARIMA HTS é uma escolha superior. O ARIMA HTS não só demonstrou melhores resultados nas métricas de avaliação, mas também é mais adequado para capturar padrões sazonais e lidar com a estrutura hierárquica dos dados. Isso o torna mais confiável para previsões em diferentes níveis, desde o país até o produto específico em cada loja.

Portanto, para a tarefa de previsão de demanda para os próximos três meses, especialmente quando se trata de identificar sazonalidades e prever com precisão em múltiplos níveis hierárquicos, o ARIMA HTS é a melhor escolha, conforme evidenciado pelas métricas de avaliação utilizadas.

7. Estimativa de Necessidade de Zíperes para Produção Global de Luvas nos Próximos Três Meses

A estimativa para a produção mundial de luvas nos próximos três meses indica que serão necessários aproximadamente 3.633 zíperes para atender à demanda projetada. Esse número foi calculado com base nas previsões de produção e nas especificações dos produtos, garantindo que a quantidade de zíperes seja suficiente para suprir a produção planejada durante o período. Essa estimativa é crucial para o planejamento de compras de matéria-prima e para assegurar que a produção ocorra sem interrupções.