Analyzing-Super-Market-Sales-Trends

This project aims to predict weekly sales for supermarkets while analyzing the factors influencing sales. This analysis will inform strategies for optimizing inventory management and resource allocation.

Implementation Approach

Create an S3 Bucket:

Go to the AWS Management Console and navigate to the S3 service.

Click on "Create bucket".

Choose a globally unique name for your bucket, select the region, and configure other settings as needed.

Create the bucket.

Set Up IAM Role:

Create an IAM role with necessary permissions to access S3 and EMR.

Attach policies like AmazonS3FullAccess and AmazonElasticMapReduceFullAccess to this role.

Launch EMR Cluster:

Go to the EMR service in the AWS Management Console.

Click on "Create cluster" and choose "Go to advanced options".

Select the EMR release, and configure hardware and software settings.

In the "Edit software settings" section, choose Spark and JupyterHub.

Choose the IAM role created in step 2.

Configure networking and security settings as per your requirements.

Launch the cluster.

Access EMR Cluster:

Once the cluster is in the "Waiting" state, you can access the EMR web interface by navigating to the provided EMR master node public DNS.

Use SSH to connect to the master node of the cluster.

Set Up EMR Studio:

Go to the EMR service in the AWS Management Console.

Click on "Create studio".

Choose the appropriate settings, including the IAM role, VPC, and subnet.

Select the EMR release and configure networking and security settings.

Create the studio.

Access EMR Studio:

Once the studio is created, you can access it from the EMR Studio dashboard in the AWS Management Console.

Launch Jupyter notebooks from within the studio environment.

Store Data in S3:

Use the AWS CLI or SDK to upload data to the S3 bucket created in step 1.

Organize your data by creating folders within the bucket, especially for storing data analyzed in Jupyter Notebook.

set Up Amazon QuickSight: Go to the Amazon QuickSight console and sign in. If you haven't used QuickSight before, you may need to set up an account.

Create a New Analysis: Once logged in, click on the "New analysis" button to start a new analysis.

Choose Your Data Source: In the data source selection screen, choose "S3" as your data source.

Connect to Your S3 Bucket: QuickSight will prompt you to connect to your S3 bucket. Provide the necessary permissions and select the bucket where your data is stored.

Select Your Data File: After connecting to your S3 bucket, navigate to the folder containing your data file and select it.

Load Your Data: QuickSight will then load your data into the analysis.

Explore Your Data: Once your data is loaded, you can start exploring it within QuickSight. You can create visualizations, filter data, and perform various types of analysis.
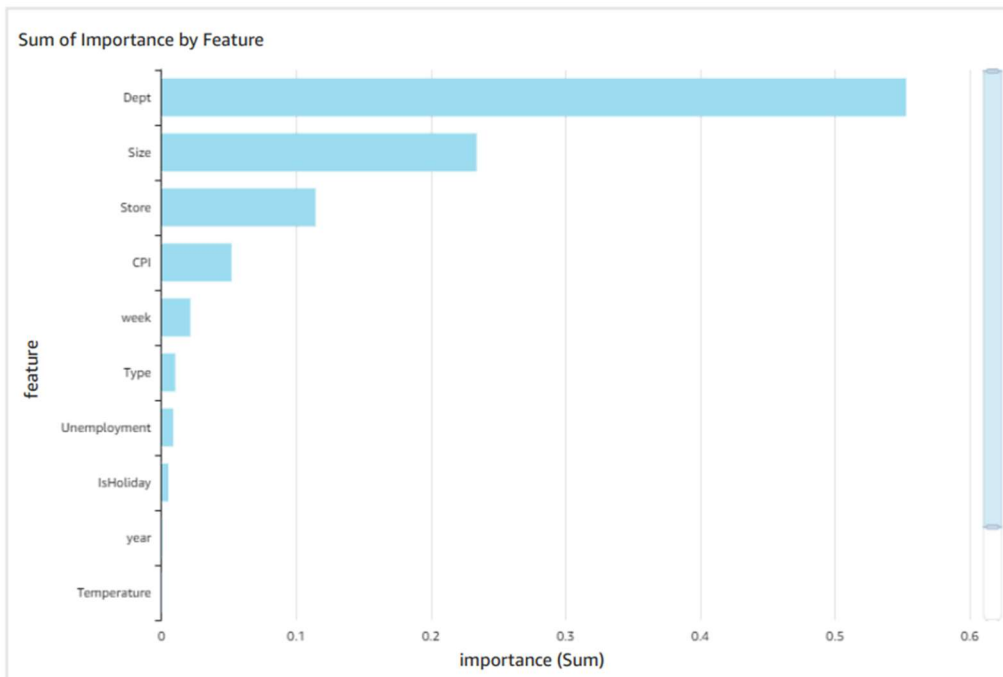
Create Visualizations: Use the QuickSight interface to create visualizations such as charts, graphs, and tables to represent your data.
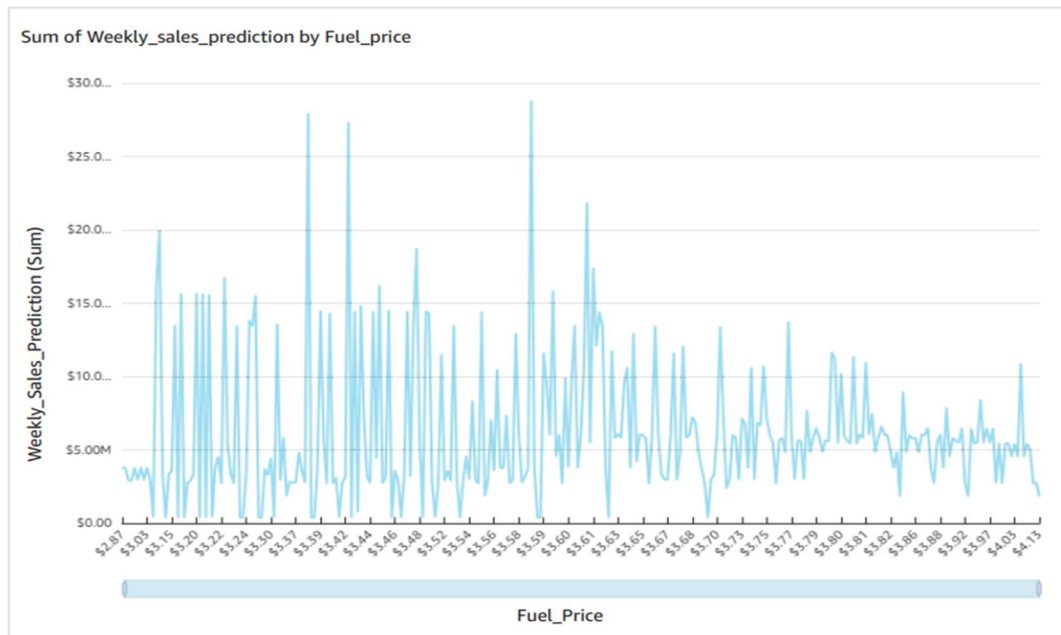
Results:

1.**Sales Predicted**
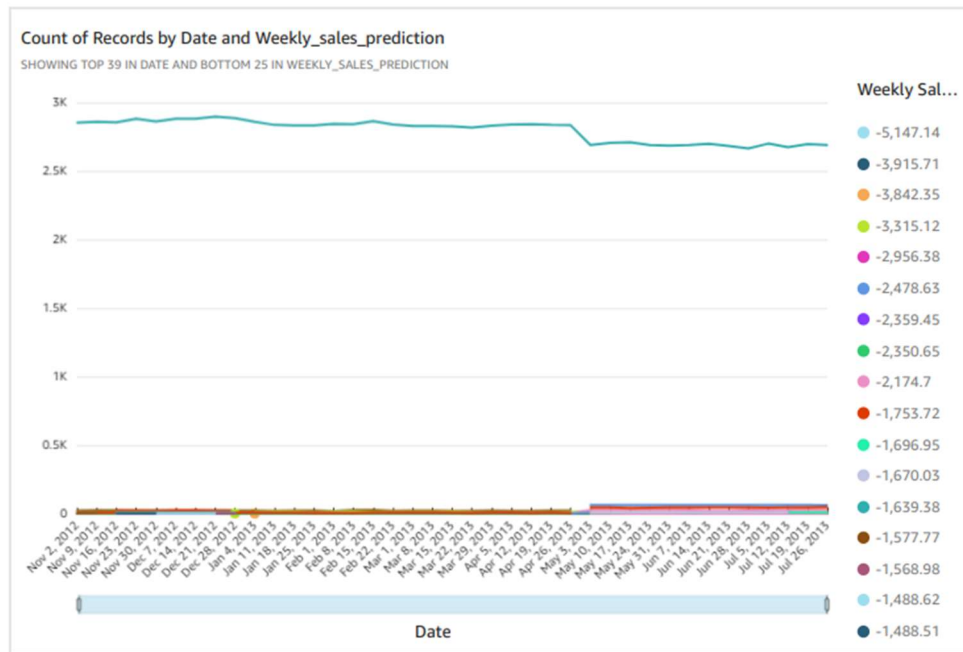
- 2. Feature importance

Fuel prices and weekly sales share a correlation, with sales demonstrating a decline during periods of higher fuel costs

Temperature and weekly sales exhibit a correlation, with sales showing lower figures during both the highest and lowest temperatures, while demonstrating higher sales during moderate temperatures.