

PARSHVANATH CHARITABLE TRUST'S
A.P. Shah Institute of Technology
Thane, 400615

Academic Year: 2022-23
Department of Computer Engineering

CSL605 SKILL BASED LAB COURSE: CLOUD COMPUTING

Mini Project Report

- **Title of Project** **Data Visualization using Amazon QuickSight**
:
- **Year and Semester** **T.E. (SEM VI)** :
- **Group Members Name and Roll No. :** Aamir Baig (07)
Manthan Agarwal(02)
Shreyas Bhalerao(12)

Table of Contents

Sr. No.	Topic	Page No.
1.	Problem Definition	3
2.	Introduction	4
3.	Description	5
4.	Implementation details	7
5.	Learning Outcome	10

Problem Definition

Ecommerce businesses rely on data to make informed decisions about their products, pricing, marketing strategies, and customer experience. This data can come from various sources such as website traffic, customer behavior, social media, and sales data. However, with the increasing volume and complexity of data, it can be challenging for ecommerce businesses to extract meaningful insights and identify actionable patterns from their data. This is where data visualization comes in.

Data visualization refers to the process of representing data graphically, in a way that makes it easier to understand and analyze. By visualizing data, ecommerce businesses can quickly identify trends, patterns, and correlations that may not be apparent in raw data. They can also explore data interactively, in real-time, and from different perspectives.

The challenge is to create a data visualization system that can effectively analyze and represent ecommerce data. Such a system should be able to handle large amounts of data and provide real-time analytics. It should also be user-friendly for non-technical users who may not have experience working with data. This means that the system should provide clear, intuitive, and customizable visualizations that allow users to explore and interpret data easily.

The visualization system should also be customizable to fit the specific needs of different ecommerce businesses. For example, some businesses may need to track customer behavior, such as click-through rates, bounce rates, and conversion rates, while others may need to monitor inventory and supply chain data. Additionally, businesses may want to optimize their marketing campaigns by tracking customer demographics, engagement rates, and purchase behavior.

Overall, the goal of ecommerce data visualization is to provide businesses with a comprehensive and intuitive data visualization system that can help them better understand their customers and improve their overall performance. By providing clear and actionable insights, data visualization can help ecommerce businesses make informed decisions, increase their revenue, and enhance customer satisfaction.

Introduction

This project uses Amazon Web Services and Amazon QuickSight service to help create data visualizations for a sample ecommerce data.

The Data used is a sample set of the amazon best-selling products, sourced from Bright Data.

Amazon QuickSight is a cloud-based business intelligence and data visualization tool that allows ecommerce businesses to quickly and easily create visualizations and reports from their data. It is a fully managed service that can connect to a wide range of data sources, including AWS data services, databases, and third-party applications.

One of the key advantages of using Amazon QuickSight is its ease of use. The platform provides a user-friendly interface that allows even non-technical users to create interactive dashboards and reports. Users can drag and drop data onto the platform and use intuitive tools to create visualizations and charts. QuickSight also provides users with a library of pre-built visualizations and templates that they can use to get started quickly.

Another advantage of QuickSight is its ability to handle large amounts of data. The platform can process and analyze data in real-time, which means that ecommerce businesses can get insights into their data quickly and make informed decisions in near real-time. QuickSight also supports advanced analytics, such as machine learning and predictive modeling, which can help ecommerce businesses identify patterns and trends in their data.

QuickSight provides a wide range of visualization options, including line charts, bar charts, scatter plots, heat maps, and more. Users can also customize their visualizations with a variety of colors, fonts, and themes to fit their branding. The platform also provides advanced features such as drill-downs, filters, and animations that allow users to explore their data in greater detail.

Collaboration and sharing are also key features of QuickSight. Users can share their dashboards and reports via email or embed them in their own websites or applications. The platform also supports user-level permissions and access controls, which means that users can control who has access to their data.

Overall, Amazon QuickSight is a powerful and user-friendly data visualization tool that can help ecommerce businesses make sense of their data and make informed decisions. Its ease of use, scalability, and flexible pricing model make it an ideal choice for businesses of all sizes.

Description

Amazon Web Services (AWS) is a cloud-based platform that provides a wide range of services for businesses, including data storage, data processing, and data analysis. One of the services provided by AWS is Amazon QuickSight, a cloud-based business intelligence and data visualization tool that allows businesses to create visualizations and reports from their data.

Ecommerce data visualization refers to the process of visually representing data from ecommerce businesses in order to gain insights, identify trends, and make informed decisions. Ecommerce businesses generate vast amounts of data, including sales data, customer data, and inventory data, among others. Data visualization allows businesses to turn this data into meaningful insights that can help drive growth, improve customer experiences, and optimize operations.

Data visualization typically involves the use of charts, graphs, and other visual elements to represent data in a way that is easy to understand and interpret. For example, a line chart may be used to visualize changes in sales over time, while a pie chart may be used to visualize the breakdown of sales by product category. Ecommerce businesses may also use heat maps, scatter plots, and other advanced visualization techniques to gain deeper insights into their data.

The goal of ecommerce data visualization is to help businesses make data-driven decisions. By visualizing their data, ecommerce businesses can identify patterns and trends that may not be immediately apparent in raw data. For example, a business may use data visualization to identify which products are selling the best, which marketing campaigns are most effective, or which regions are generating the most sales.

Data visualization also allows ecommerce businesses to communicate insights and information to stakeholders in a clear and concise way. For example, a business may create a dashboard that displays key performance indicators (KPIs) such as sales, conversion rates, and customer retention, allowing stakeholders to quickly and easily understand the business's performance.

Overall, ecommerce data visualization is a powerful tool that can help businesses turn their data into actionable insights. By using visualization techniques to identify patterns and trends, ecommerce businesses can make informed decisions that can help drive growth and improve customer experiences.

The project aims to create data visualizations for a sample ecommerce data set of the Amazon best-selling products using AWS and QuickSight. The data set includes information on the products, such as the product name, product category, price, and sales rank.

The first step in the project would be to import the sample data set into AWS. AWS offers a variety of data storage services, such as Amazon S3, Amazon RDS, and Amazon DynamoDB. Depending on the structure and size of the data set, the appropriate storage service can be chosen. We chose S3 as it is adequate for the sample data.

Once the data set is stored in AWS, it can be processed and analyzed using AWS data processing services such as Amazon EMR, Amazon Athena, or Amazon Redshift. These services can be used to transform the raw data into a format suitable for data visualization.

The next step would be to connect the data set to Amazon QuickSight. QuickSight offers several options for connecting to data sources, including AWS data services, databases, and third-party applications. Once the data set is connected, QuickSight can be used to create visualizations and reports.

We used a sample dataset from BrightData for our project.

QuickSight provides a wide range of visualization options, including line charts, bar charts, scatter plots, heat maps, and more. Users can customize their visualizations with a variety of colors, fonts, and themes to fit their branding. The platform also provides advanced features such as drill-downs, filters, and animations that allow users to explore their data in greater detail.

The project aims to create a dashboard that displays key performance indicators (KPIs) such as sales, revenue, and product category. The dashboard would allow stakeholders to quickly and easily understand the business's performance and make data-driven decisions.

Overall, the project aims to demonstrate the power of AWS and QuickSight in creating data visualizations for ecommerce businesses. By using AWS to store and process the data and QuickSight to create visualizations, ecommerce businesses can gain valuable insights into their data and make informed decisions.

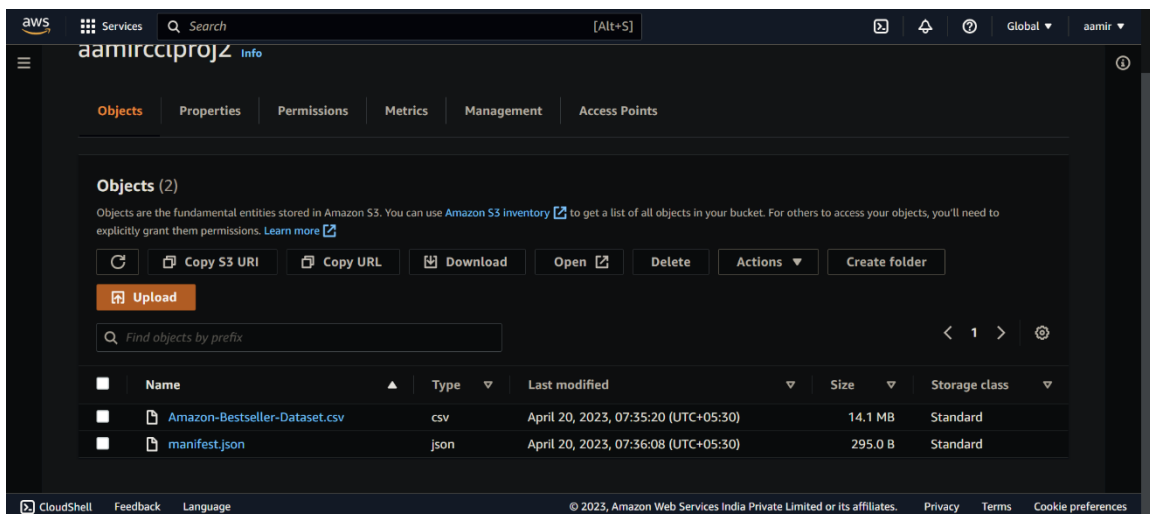
Implementation Details

A new S3 bucket was created to host the data on AWS. This bucket contained the amazon best seller dataset and a manifest.json file to help quicksight access the data from the bucket.

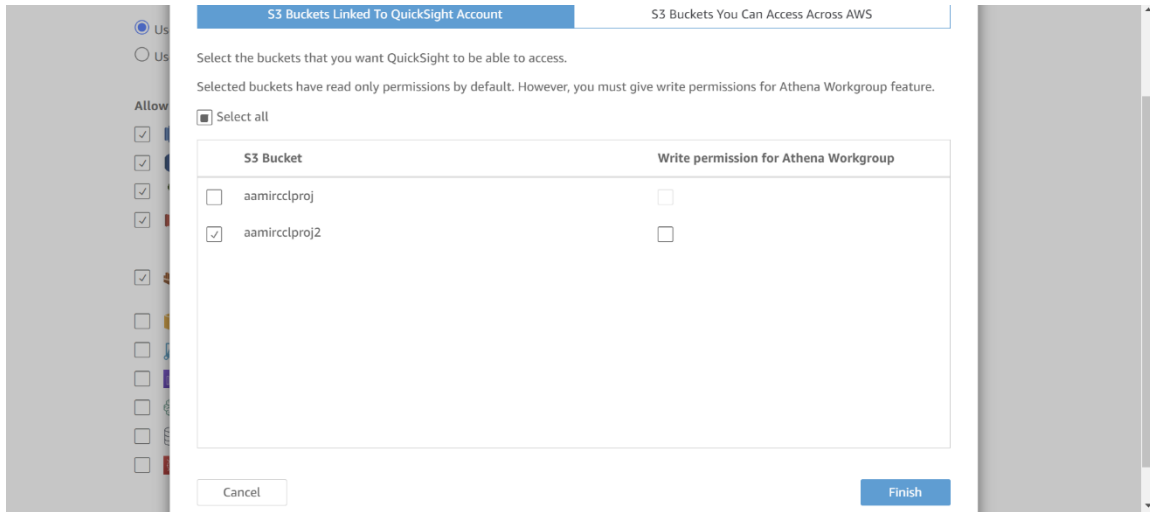
The manifest.json file

```
1  {
2    "fileLocations": [
3      {
4        "URIs": [
5          "s3://aamircclproj2/Amazon-Bestseller-Dataset.csv"
6        ]
7      }
8    ],
9    "globalUploadSettings": {
10     "format": "CSV",
11     "delimiter": ",",
12     "textqualifier": "\"",
13     "containsHeader": "true"
14   }
15 }
16
```

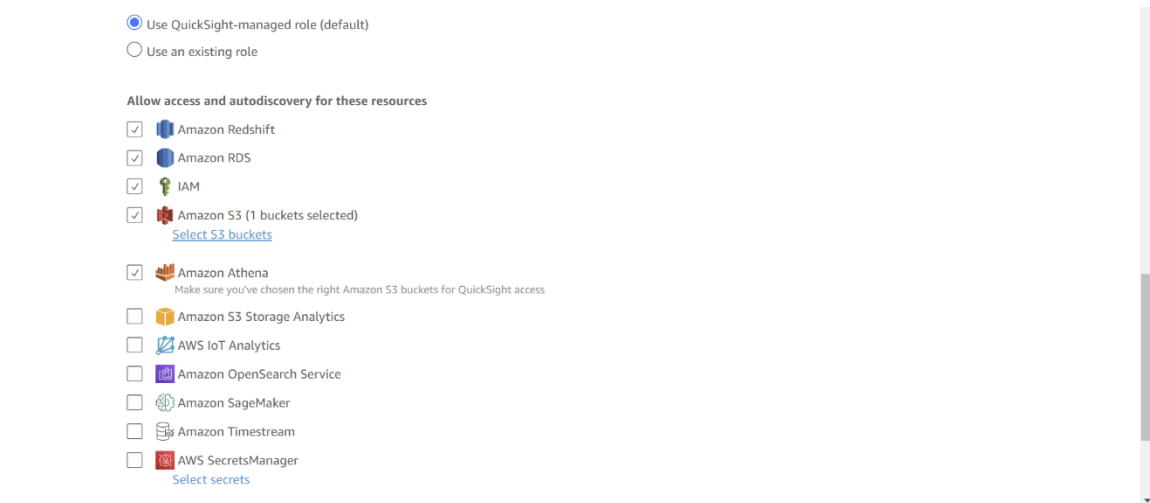
Uploading dataset and manifest file to the S3 bucket.



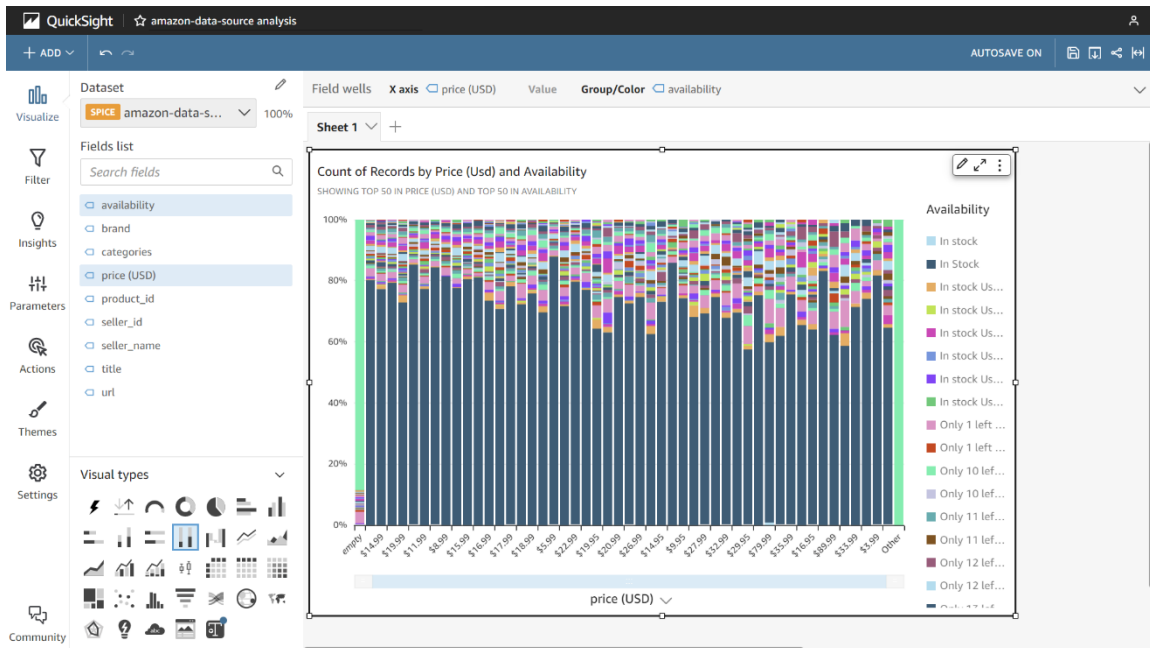
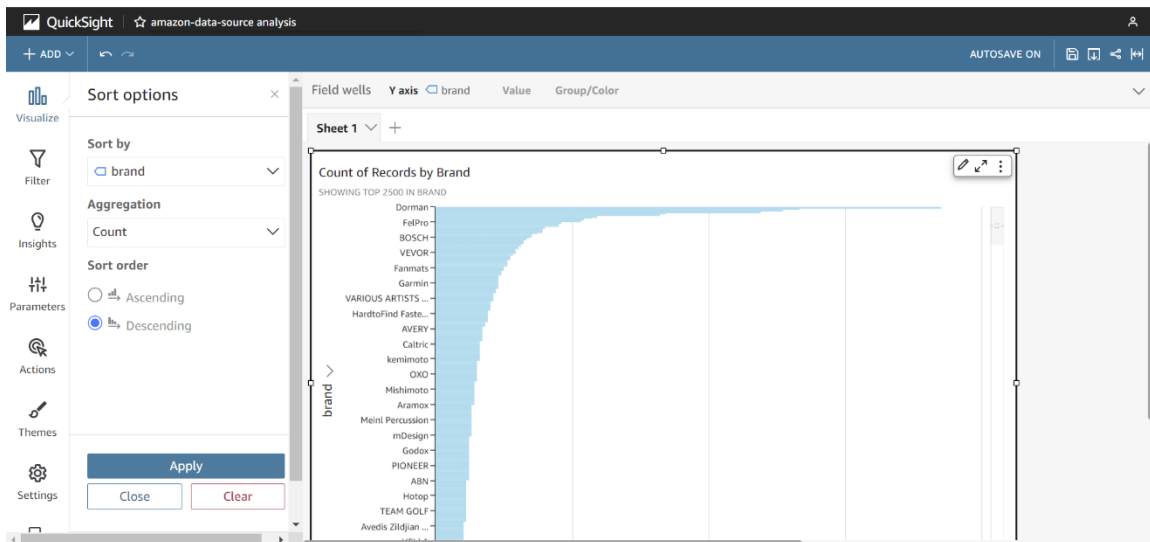
Selecting the S3 bucket source in QuickSight



Selecting the S3 bucket module



Visualizing data using QuickSight



Learning Outcome

- Learned to import a data set into AWS S3 and set up appropriate access controls to ensure the security of the data.
- Learned interfacing and connecting data sources to Amazon QuickSight and created various visualizations and reports.
- Created a dashboard in Amazon QuickSight that displayed key performance indicators (KPIs) such as pricing, revenue, and product category.
- Gained experience in customizing visualizations with a variety of colors, fonts, and themes and axis range manipulation.
- Learned how to effectively communicate insights and information to stakeholders through clear and concise data visualizations.
- Gained a deeper understanding of how AWS and QuickSight can be used to create data visualizations for ecommerce businesses and make informed decisions based on data insights.
- Became comfortable in working with cloud-based tools and services and understanding their advantages over on-premises solutions.