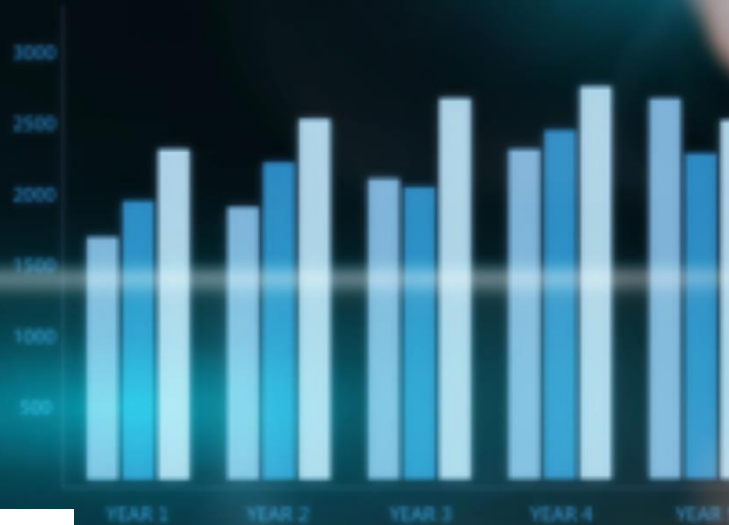




QuickSight



Software Evaluation

Binge Bro's

Rohit Shivaji Bhoite

Nitish Halaganahalli Reddy

Sravya Buridi

Tejaswini Sapkota

Introduction

Amazon QuickSight was launched in November 2016 by Amazon Web Services (AWS) as a cloud-based business intelligence (BI) service. Its primary aim was to provide a cost-effective and user-friendly way for organizations to extract insights from their data. Initially, QuickSight was designed to connect to data sources in the AWS cloud, but it has since expanded to include data from a range of sources, including third-party applications and on-premises databases.

In 2017, AWS added support for machine learning capabilities, enabling users to incorporate pre-built or custom machine learning models into their dashboards. The following year, AWS introduced AutoGraph, a feature that uses machine learning algorithms to suggest the most appropriate visualizations for a given dataset. Since then, AWS has continued to develop and improve QuickSight, adding new features and functionality, such as collaboration capabilities, expanded visualization options, and integration with other AWS services like SageMaker. As a fully managed cloud-based service, QuickSight offers enterprise-grade security, global availability, and built-in redundancy, making it a reliable and secure option for organizations. With its ability to generate interactive dashboards and visualizations, incorporate machine learning models, and handle large volumes of data, QuickSight is a valuable asset for businesses in various industries.

Our project will utilize QuickSight and a PostgreSQL database to demonstrate the use of Business Intelligence & Analytics in a real-world application. We aim to showcase the capabilities of QuickSight in creating interactive visualizations and the integration of a database to store and retrieve data. By using these tools, we hope to demonstrate the importance of BI&A in decision making and how it can be used to gain insights from data. Additionally, we will also be highlighting the benefits of using such technologies in streamlining and automating the data analysis process. Overall, the project aims to provide a hands-on understanding of the practical applications of BI&A and the role of technologies like QuickSight and PostgreSQL in implementing them.

USE CASE

Several companies, including PeopleScout, CapitalOne, and the NFL, use Amazon QuickSight to gain insights and make data-driven business decisions. A manager at PeopleScout can automate basic reporting and dashboarding and deliver results through QuickSight while an employee at CapitalOne visualizes financial data, detects fraud, analyzes customer behavior, and optimizes marketing campaigns using Amazon QuickSight. Meanwhile, the employees at NFL analyze player performance, fan engagement, and revenue and business metrics to make informed decisions using QuickSight.

AWS QuickSight is used in several industries and below are examples of they can be utilized

Healthcare industry: Healthcare providers can use QuickSight to analyze patient data and identify patterns that can help in early diagnosis and treatment of diseases. They can also use the tool to track and analyze operational data such as patient wait times and resource utilization.

Financial services industry: QuickSight can be used by banks and financial institutions to analyze customer data and identify trends that can help in fraud detection and risk management. They can also use the tool to track and analyze market data and make informed investment decisions.

Manufacturing industry: Manufacturing companies can use QuickSight to analyze production data and identify areas where they can optimize their processes. They can also use the tool to track and analyze supply chain data to identify potential bottlenecks and make informed decisions.

Education industry: Educational institutions can use QuickSight to analyze student performance data and identify areas where they can improve their teaching methodologies.

With its powerful features and ability to integrate with various data sources, QuickSight has become a preferred choice for many companies looking to improve their business performance.

Software Details

AWS QuickSight is a business intelligence (BI) service provided by Amazon Web Services (AWS) that enables customers to create interactive visualizations, perform ad-hoc analysis, and share insights with others in their organization. As a cloud-based service, it offers a scalable and cost-effective solution for analyzing and visualizing data.

Here are the main components of Amazon QuickSight:

Data Sources: QuickSight integrates with a wide range of data sources, including AWS services such as S3, RDS, Redshift, and Athena, as well as third-party services like Salesforce, MySQL, and PostgreSQL. It is important to determine which data sources are relevant to your organization and ensure they are properly configured for use with QuickSight.

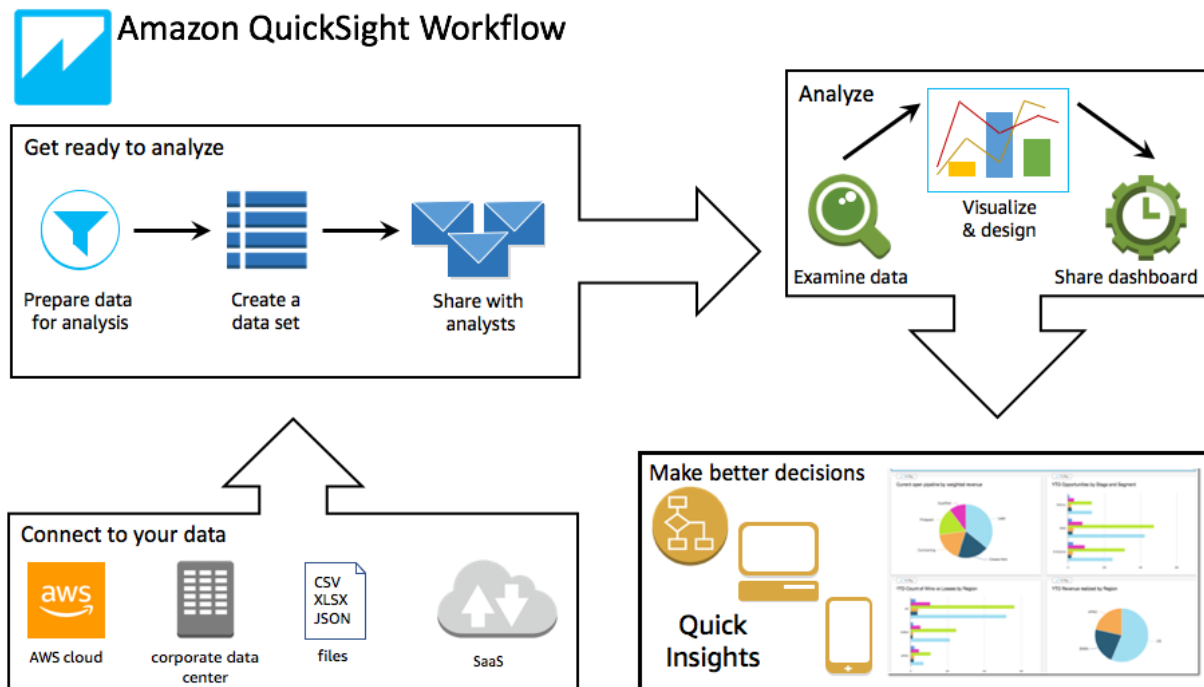
Security: Security is a critical concern when integrating QuickSight into your BI and IT architecture. AWS provides a range of security features, such as encryption, access controls, and audit logs, to help ensure the confidentiality and integrity of your data. It is important to review and configure these security features appropriately to meet your organization's security and compliance requirements.

Analysis: QuickSight allows users to create visualizations, charts, and graphs to analyze their data. It includes a wide range of visualization options such as bar charts, line charts, heat maps, pivot tables, and more.

Dashboard: Users can create interactive dashboards that can be shared with other users or embedded into other applications. Dashboards can include multiple visualizations and can be customized to suit the user's needs.

Administration: QuickSight includes features for administration and management, such as user access control, data source management, and audit logging

Overall, Amazon QuickSight's components work together to provide a comprehensive business intelligence and data visualization solution for businesses of all sizes.



1. We must create a new analysis first.
2. We have the option to add new or existing database.
3. For chart creation we must choose fields. QuickSight automatically suggests the best visualization.
4. There is option to add multiple charts, tables, or insights to the analysis. Can resize and rearrange them on one or more sheets. Use extended features to add variables, custom controls, colors, additional pages (called sheets), and more.
5. Dashboard can be created and published to share it with other people.

KEY ALTERNATIVES

KEY COMPETITORS AND PRODUCTS

Some of the key competitors and products in this space include:

Microsoft Power BI: Microsoft Power BI is a business analytics service that enables interactive visualizations, data analytics, and business intelligence. The service provides a variety of data connectors and integrates with multiple Microsoft products.

Tableau: Tableau is a business intelligence software that emphasizes data visualization and analysis, providing a variety of tools for both. It offers drag-and-drop functionality and is known for its strong focus on visualizing data.




Google Data Studio: Google Data Studio is a cloud-based data visualization tool that empowers users to create personalized reports and dashboards. The tool provides numerous data connectors and integrates with various Google products.

QlikView: QlikView is a data discovery and business intelligence platform that enables interactive data visualization and analysis. The platform offers data connectors and integrates with various third-party data sources.

Zoho Analytics: Zoho Analytics is a cloud-based business intelligence and analytics platform that provides a range of features and capabilities to help businesses and organizations analyze their data and gain insights into their operations.

Infogram: Infogram is a cloud-based tool for data visualization that enables users to create and share interactive charts, graphs, maps, and infographics.

2. STRENGTHS / WEAKNESSES TO COMPETITORS

Tool	Strengths	Weakness
	<p>Powerful data visualization capabilities.</p> <p>Integration with other Microsoft tools such as Excel, SharePoint, and Teams.</p> <p>Customizable dashboards and reports.</p> <p>Strong security and compliance features.</p>	<p>Steep learning curve for beginners.</p> <p>Limited data preparation capabilities.</p> <p>Some advanced features require a paid Pro version.</p>
	<p>Highly intuitive and user-friendly interface.</p> <p>Ability to handle large and complex datasets.</p> <p>Comprehensive set of data visualization options.</p> <p>Offers both cloud-based and on-premises deployment options.</p>	<p>Expensive licensing costs for both desktop and server versions.</p> <p>Requires significant IT resources to set up and maintain.</p> <p>Limited capabilities for data preparation and cleansing.</p>
	<p>Completely free to use.</p> <p>Easy integration with other Google tools such as Google Analytics and Google Sheets.</p> <p>Simple and intuitive interface.</p> <p>Customizable dashboards and reports.</p>	<p>Limited data visualization options.</p> <p>Limited data preparation capabilities.</p> <p>Limited collaboration features compared to other BI tools.</p>

	<p>Strong data visualization capabilities.</p> <p>Ability to handle complex data models.</p> <p>Advanced capabilities for data transformation and cleansing.</p> <p>Offers both cloud-based and on-premises deployment options.</p>	<p>Expensive licensing costs.</p> <p>Steep learning curve for beginners.</p> <p>Limited collaboration features.</p>
	<p>Ease of Use</p> <p>Integration with Zoho Suite</p> <p>Customizable Dashboards</p> <p>Data Blending</p> <p>Share reports and dashboards with other users and collaborate on data analysis.</p>	<p>Limited Customization</p> <p>Data Refresh Limits</p> <p>Steep Learning Curve for Advanced Features</p> <p>Limited Integration with Third-Party Tools</p>
	<p>User-friendly Interface</p> <p>Wide Range of Templates and Visuals</p> <p>Interactive Features</p> <p>Collaboration Features</p> <p>Integration with Other Tools</p>	<p>Limited Customization</p> <p>Limited Data Import and Manipulation</p> <p>Limited Export Options</p> <p>Cost</p>

STRENGTHS AND WEAKNESS OF QUICKSIGHT

Strengths:

Scalability: QuickSight is a fully managed, cloud-based service that is highly scalable. It can handle large volumes of data and support thousands of users without requiring any hardware or software upgrades.

Easy to use: QuickSight's intuitive drag-and-drop interface makes it easy for non-technical users to create interactive dashboards and visualizations without any coding or IT support.

Integration with AWS: QuickSight integrates seamlessly with other AWS services such as Amazon S3, Amazon Redshift, and Amazon RDS, making it easy to access and analyze data stored in these services.

Cost-effective: QuickSight offers a pay-per-session pricing model, which means that you only pay for the number of user sessions that you use. This can be cost-effective for organizations with variable user loads.

Mobile support: QuickSight offers mobile support, allowing users to access their dashboards and visualizations from mobile devices.

Weakness:

Limited data connectivity: QuickSight offers connectivity to a limited number of data sources, which may be a disadvantage for organizations with data stored in multiple platforms.

Limited customization: QuickSight offers a limited number of customization options, which may limit the flexibility of visualizations and dashboards.

Limited reporting capabilities: QuickSight offers limited reporting capabilities, which may be a disadvantage for organizations that require advanced reporting features.

Limited data preparation: QuickSight's data preparation features are relatively limited compared to other BI services, which may require organizations to use third-party data preparation tools.

Limited collaboration: QuickSight's collaboration features are relatively limited compared to other BI services, which may be a disadvantage for organizations that require extensive collaboration and sharing capabilities.

SOFTWARE RECOMMENDATION

Amazon QuickSight provides various benefits, such as its intuitive interface, affordability, and scalability. Its user-friendly design allows non-technical users to create interactive visualizations and dashboards easily. Additionally, QuickSight employs a pay-as-you-go pricing model, which eliminates the need for upfront investments in hardware or software. Furthermore, QuickSight is a highly scalable solution capable of handling large volumes of data, making it an ideal choice for organizations with diverse data needs.

However, to further enhance its value, there are a few key areas that can be improved upon. These include enhancing collaboration features to allow for simultaneous work on the same dashboard, expanding visualization options to include new chart types and more customization options, and increasing integration with other AWS services such as RDS, SageMaker. Finally, improving data preparation capabilities within the tool could save time and effort for analysts by reducing the need to clean and format data outside of QuickSight. By addressing these areas of improvement, Amazon QuickSight can become an even more valuable tool for organizations seeking to gain insights from their data.

FUTURE DEVELOPMENTS

QuickSight now offers Machine Learning (ML) Insights, which is a new feature that enables users to discover hidden trends and insights in their data with the help of machine learning algorithms. This feature includes predictive analytics, anomaly detection, and natural language generation (NLG) to help users better understand their data.

Recently, QuickSight has introduced Embedded Analytics, a feature that enables developers to seamlessly integrate analytics dashboards and reports into their applications. This functionality simplifies the process of delivering data-based insights to users, eliminating the need for them to switch between different applications.

QuickSight has introduced customizable themes, giving users the ability to design and apply their own branded themes to their dashboards and reports. Users can personalize various design elements such as colors, fonts, and other visual components to align with their organization's branding.

It has recently updated its offerings with new and enhanced visualizations, such as Box Plots, Funnel Charts, and Heat Maps, to assist users in comprehending their data and recognizing trends and patterns. To optimize data analysis, we can leverage Machine Learning (ML) Insights and upgraded visualizations in our projects.

CONCLUSIONS

Amazon QuickSight is a powerful data visualization and business intelligence tool that offers several advantages for users. With its ease of use, cost-effectiveness, and scalability, QuickSight is a popular choice for organizations looking to gain insights from their data. However, there are still some areas where QuickSight can be improved upon, such as enhancing collaboration features, expanding visualization options, improving integration with other AWS services, and enhancing data preparation capabilities.

Moreover, QuickSight's integration with other AWS services such as SageMaker, S3, RDS and Redshift allows users to easily access and analyze their data without the need for complex data migration or integration processes. QuickSight also offers a range of sharing options, enabling users to share their visualizations and dashboards securely with colleagues and clients. The pay-as-you-go pricing model means that users only pay for what they use, reducing the financial burden on organizations.

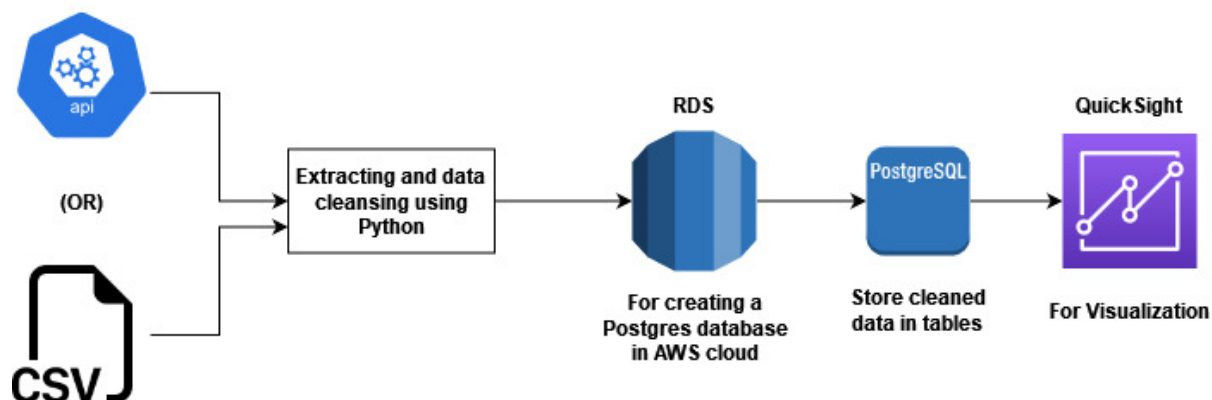
QuickSight's user-friendly interface, along with its ability to handle various data sources and formats, makes it a versatile and flexible tool for businesses of all sizes. Furthermore, the tool's built-in redundancy, global availability, and enterprise-grade security ensure that users can trust QuickSight with their sensitive business data. As QuickSight continues to evolve, it has the potential to become an even more valuable tool for organizations seeking to make data-driven decisions. With its ability to handle large volumes of data, generate interactive dashboards and visualizations, and incorporate machine learning models, QuickSight is a powerful asset for businesses in a variety of industries.

REFERENCE

1. <https://www.kaggle.com/datasets/karkavelrajaj/amazon-sales-dataset> - Datasource
2. <https://hevodata.com/learn/amazon-quicksight-a-comprehensive-guide/>
3. <https://sparkbyexamples.com/pandas/pandas-write-dataframe-to-csv-file/>
4. <https://towardsdatascience.com/amazon-rds-step-by-step-guide-14f9f3087d28>
5. <https://aws.amazon.com/blogs/big-data/transform-data-and-create-dashboards-simply-using-aws-glue-databrew-and-amazon-quicksight/>
6. <https://docs.aws.amazon.com/quicksight/latest/user/how-quicksight-works.html>
7. <https://aws.amazon.com/blogs/architecture/category/analytics/amazon-quicksight/>
8. <https://alternativeto.net/software/power-bi-for-office-365/?license=opensource>

TUTORIAL

Project Architecture



1. HOW TO GET THE SOFTWARE AND DATA

We are extracting the data in the CSV format, the data obtained is not clean and can't be used for visualization. In the ETL(Extract, Transform and Load) process, this is the Extraction phase.

File

Home

Insert

Draw

Page Layout

Formulas

Data

Review

View

Automate

Help

Load Test

↶

↷

Undo

Paste

Clipboard

Calibri

11

A[^]

A_v

Font

Alignment

Wrap Text

Merge & Center

Number

A1

:

product_id

	A	B	C	D	E	F	G
1	product_id	product_name	category	discounted_price	actual_price	discount_percent	rating
2	B07JW9H4J1	Wayona Nylon Braided U	Computers&Acces	â,1399	â,1,099	64%	4.2
3	B098NS6PVG	Ambrane Unbreakable 60	Computers&Acces	â,1199	â,1349	43%	4
4	B096MSW6CT	Sounce Fast Phone Charg	Computers&Acces	â,1199	â,1,899	90%	3.9
5	B08HDJ86NZ	boAt Deuce USB 300 2 in	Computers&Acces	â,1329	â,1699	53%	4.2
6	B08CF3B7N1	Portronics Konnect L 1.2M	Computers&Acces	â,1154	â,1399	61%	4.2
7	B08Y1TFSP6	pTron Solero TB301 3A T	Computers&Acces	â,1149	â,1,000	85%	3.9
8	B08WRWPM22	boAt Micro USB 55 Tangl	Computers&Acces	â,1176.63	â,1499	65%	4.1
9	B08DDRGWTJ	MI Usb Type-C Cable Sma	Computers&Acces	â,1229	â,1299	23%	4.3

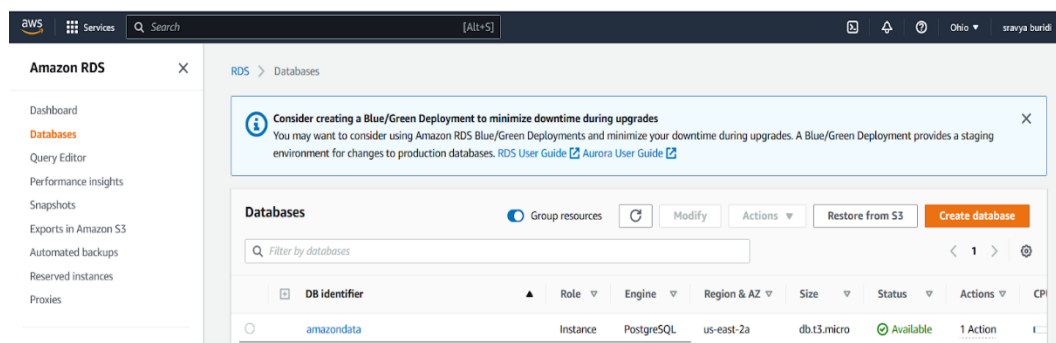
To use this data for visualization, we had to perform some data cleansing and data transformation. There are multiple ways to achieve this, we have opted to perform these tasks using Python. In the Jupyter notebook we extracted the data using 'Pandas' library and saved it as a dataframe to perform data cleansing and data transformation as required. This is the Transformation phase in the ETL process.

```
In [97]: df = pd.read_csv(r"C:\Users\nitis\OneDrive\Desktop\amazon.csv")
df
```

Out[97]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Cha...	Computers&Accessories/Accessories&Peripherals...	₹399	₹1,099	64%	4.2	24,269
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging 1.5...	Computers&Accessories/Accessories&Peripherals...	₹199	₹349	43%	4.0	43,994
2	B096MSW6CT	Source Fast Phone Charging Cable & Data Sync U...	Computers&Accessories/Accessories&Peripherals...	₹199	₹1,899	90%	3.9	7,928
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories/Accessories&Peripherals...	₹329	₹699	53%	4.2	94,363
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories/Accessories&Peripherals...	₹154	₹399	61%	4.2	16,905

To save the cleaned data, we used AWS RDS to create a database. We used AWS as it is very easily compatible with QuickSight. We found some difficulties in connecting to QuickSight from our local database in PostgreSQL. There were issues with connecting to a hostname in a local database, which was a private database.



Selecting PostgreSQL as the database

The screenshot shows the 'Engine options' section in the AWS RDS console. Under 'Engine type', several options are listed: Aurora (MySQL Compatible), Aurora (PostgreSQL Compatible), MySQL, MariaDB, PostgreSQL (selected), Oracle, and Microsoft SQL Server. The PostgreSQL option is highlighted with a blue border and a blue dot. To the right, a 'PostgreSQL' information panel is visible, providing details about the database system and its features.

Engine options

Engine type [Info](#)

- ☐ Aurora (MySQL Compatible)
- ☐ Aurora (PostgreSQL Compatible)
- ☐ MySQL
- ☐ MariaDB
- ☒ PostgreSQL
- ☐ Oracle
- ☐ Microsoft SQL Server

PostgreSQL

PostgreSQL is a powerful, open-source object-relational database system with a strong reputation of reliability, stability, and correctness.

- High reliability and stability in a variety of workloads.
- Advanced features to perform in high-volume environments.
- Vibrant open-source community that releases new features multiple times per year.
- Supports multiple extensions that add even more functionality to the database.
- Supports up to 15 Read Replicas per instance, within a single Region or 5 read replicas cross-region.
- Supports General Purpose, Memory Optimized, and Burstable Performance instance classes.

We had to create a security group using AWS EC2 to make the RDS database accessible in QuickSight.

The screenshot shows the AWS Management Console interface for the 'Security Groups' page. The page title is 'Security Groups (1/1) Info'. There is a search bar with the text 'Filter security groups' and a 'search: sg-043fdea1ba46b087e' filter applied. A table lists the security groups, with one group selected. The table has columns for Name, Security group ID, Security group name, VPC ID, Description, and Owner. The selected group is 'default' with ID 'sg-043fdea1ba46b087e' and VPC ID 'vpc-097c64523021dc190'.

Security Groups (1/1) Info

Filter security groups

search: sg-043fdea1ba46b087e

<input checked="" type="checkbox"/>	Name	Security group ID	Security group name	VPC ID	Description	Owner
<input checked="" type="checkbox"/>	-	sg-043fdea1ba46b087e	default	vpc-097c64523021dc190	default VPC security gr...	545724729218

2. HOW TO LOAD DATA

Before loading the data into the database several data transformations had to be performed using Python script, which resulted in cleaner data that is more suitable for visualization.

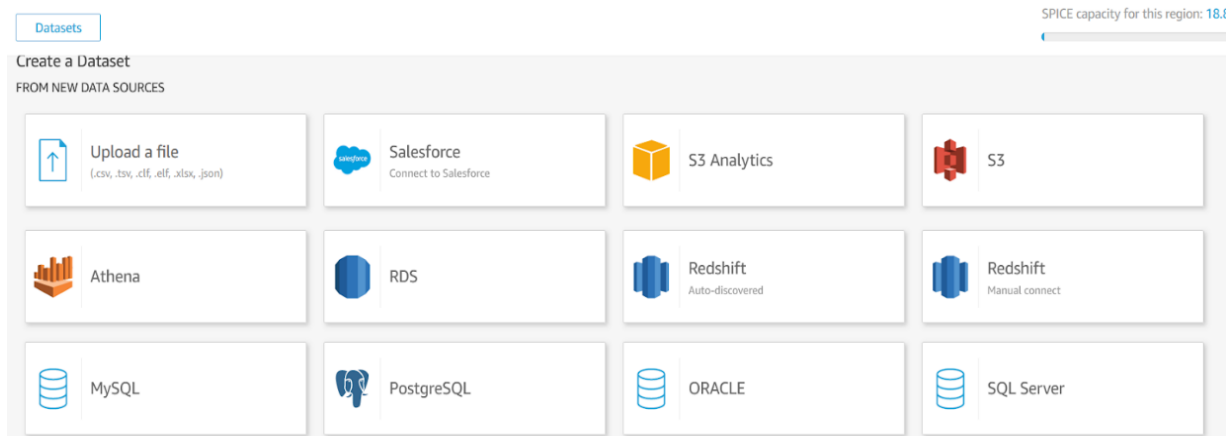
```
In [153]: data.head()
```

Out[153]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Cha...	Computers&Accessories Accessories&Peripherals ...	399.0	1099.0	0.64	4.2	24269
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging 1.5...	Computers&Accessories Accessories&Peripherals ...	199.0	349.0	0.43	4.0	43994
2	B096MSW6CT	Source Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	199.0	1899.0	0.90	3.9	7928
3	B08HJDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	329.0	699.0	0.53	4.2	94363
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	154.0	399.0	0.61	4.2	16905

To load the cleaned data table to a database using Python, certain libraries like pycpg2, pandas are used to make a connection with the database. This is the Loading phase in the ETL process.

Now with the data stored in the form of tables in the PostgreSQL database in RDS, it is easier to make a connection with QuickSight to start visualization.



By giving these credentials, we can now access tables stored in the PostgreSQL database and perform visualization.

New PostgreSQL data source ×

Data source name

Connection type

Public network ▼

Database server

Port

Database name

Username

Password

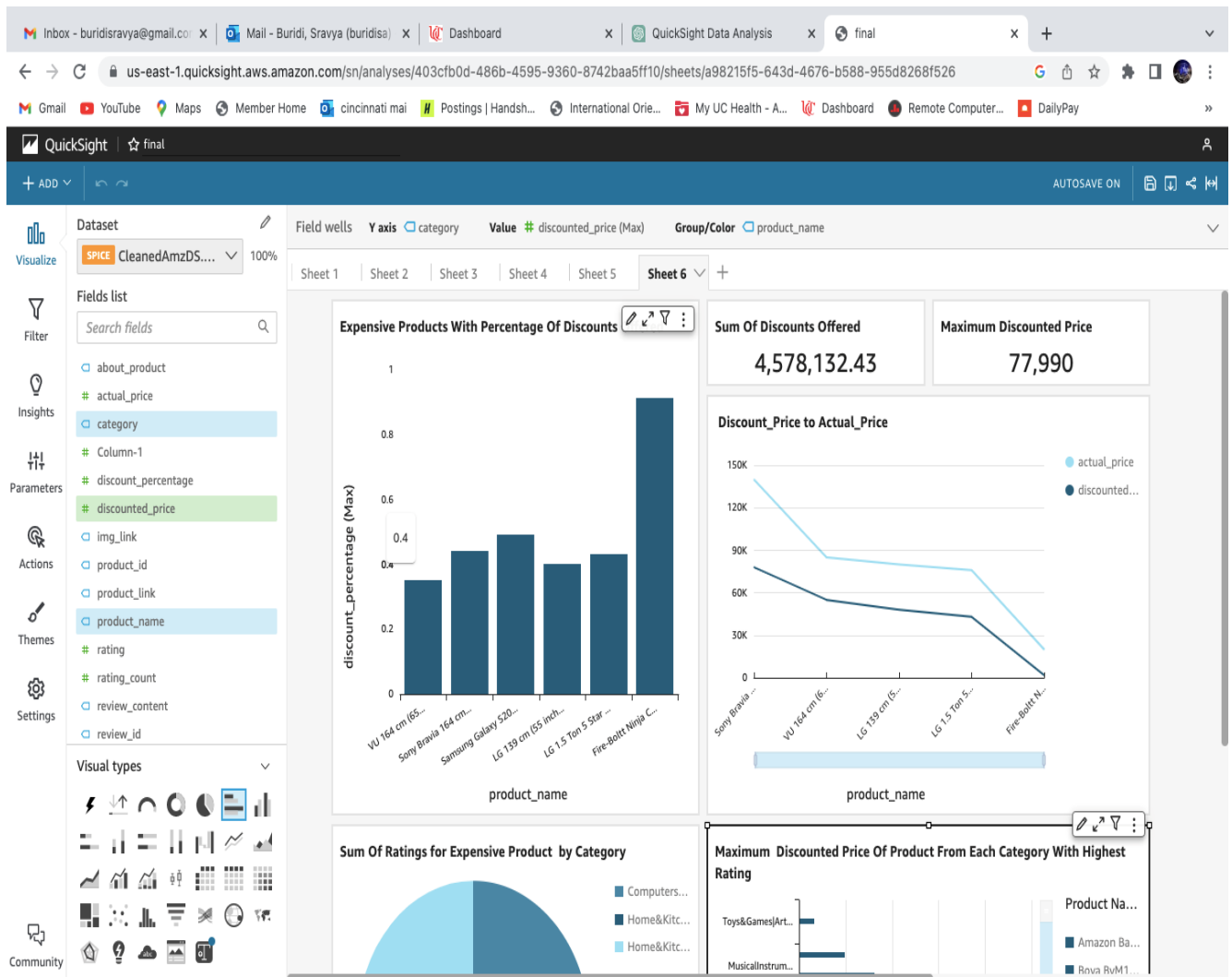
☒ Enable SSL

3. USE THE SOFTWARE TO VISUALIZE THE DATA

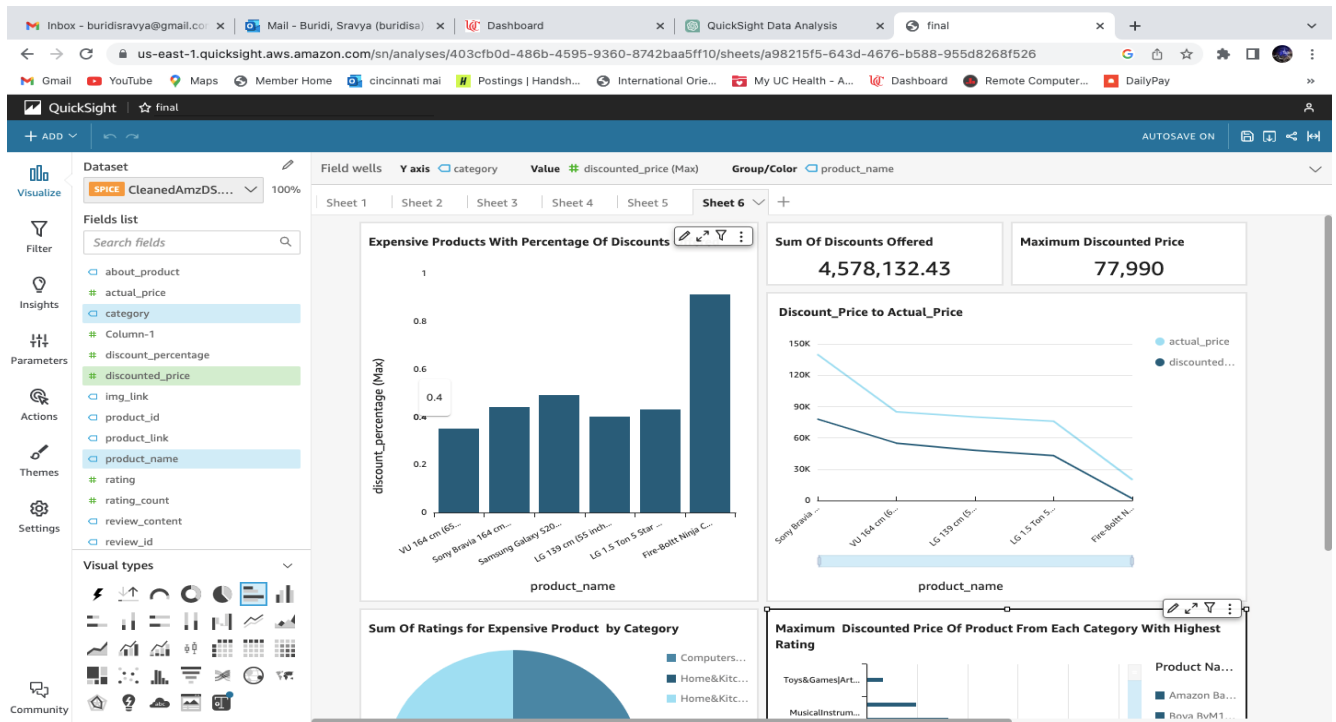
After loading the data, drag and drop interface helps us to create visualizations that help us understand our data.

In Detail:

When data is loaded to QuickSight the columns in the data are listed in fields. Select the required fields to analyze.

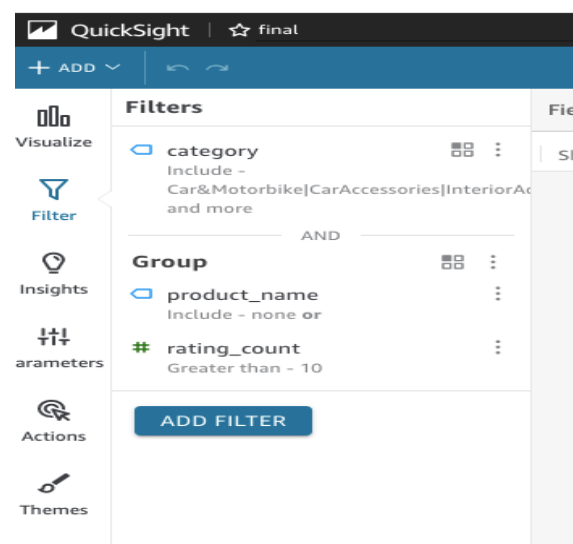
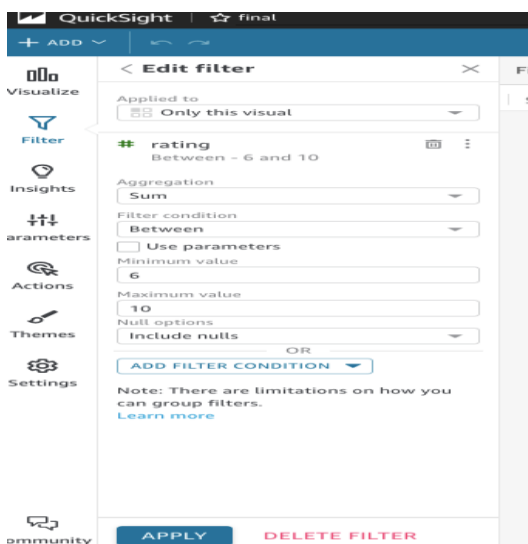


Quick Sight helps in visualizing the data through a wide variety of chart types (bar, pie, line, clustered bar, scatter, etc.) that makes it more understandable and easier to analyze.



Use dashboard filters to minimize the amount of data displayed on the graph.

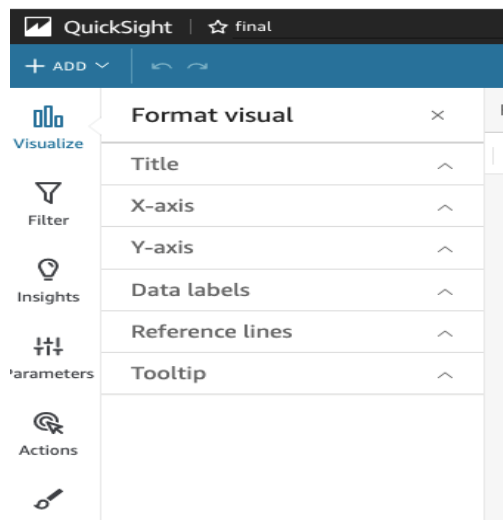
Filters can be customized by including more than one column or by selecting specific values, data ranges or numerical ranges for each field.



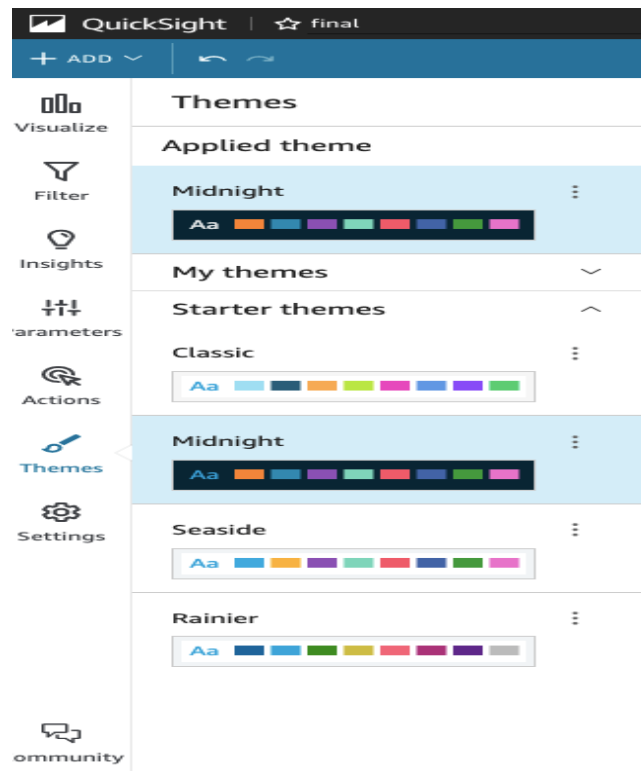
Also, QuickSight provides valuable insights by suggesting various plot types based on the data categories and displaying relevant product descriptions based on the filters applied. By doing so, QuickSight enables users to analyze their data and gain meaningful insights quickly and easily.



QuickSight enables users to format visuals to meet their specific needs. Users can customize visualizations with a range of formatting options, such as adding titles, legends, axes, and data labels. They can also adjust the size, color, and style of visual elements to suit their preferences. By formatting visuals in QuickSight, users can improve the clarity and accuracy of their data representations and better communicate their insights to others.



Using themes in QuickSight can enhance the appearance of visualizations, providing a better user experience. By applying custom styles and branding, users can select specific design elements such as colors and fonts that align with their brand or style preferences and apply them consistently across their dashboards and reports. This makes the visualizations more visually appealing and easier to interpret, ultimately enhancing the overall effectiveness of data analysis.



To export or download visuals in QuickSight, users can create the necessary visuals and place them on a single dashboard. Once the visuals are on the dashboard, users can export or download the dashboard as needed. This allows users to easily share their data and insights with others, whether it's for collaboration or reporting purposes. By centralizing the visuals on a dashboard, users can also keep their data organized and easily accessible, improving the efficiency of their data analysis processes.

