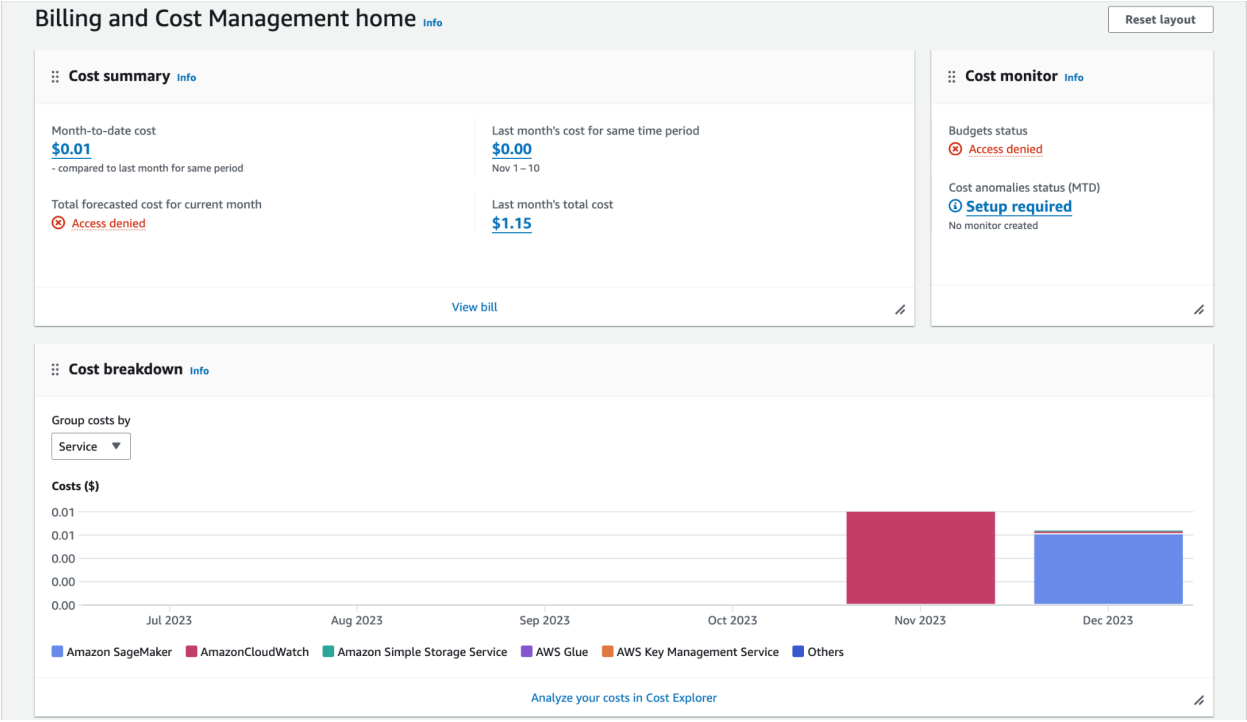
PHASE 4 Optimization, Scalability and Deployment
AWS Cost Analysis and Optimization:

Analyze the cost of running your AWS resources.

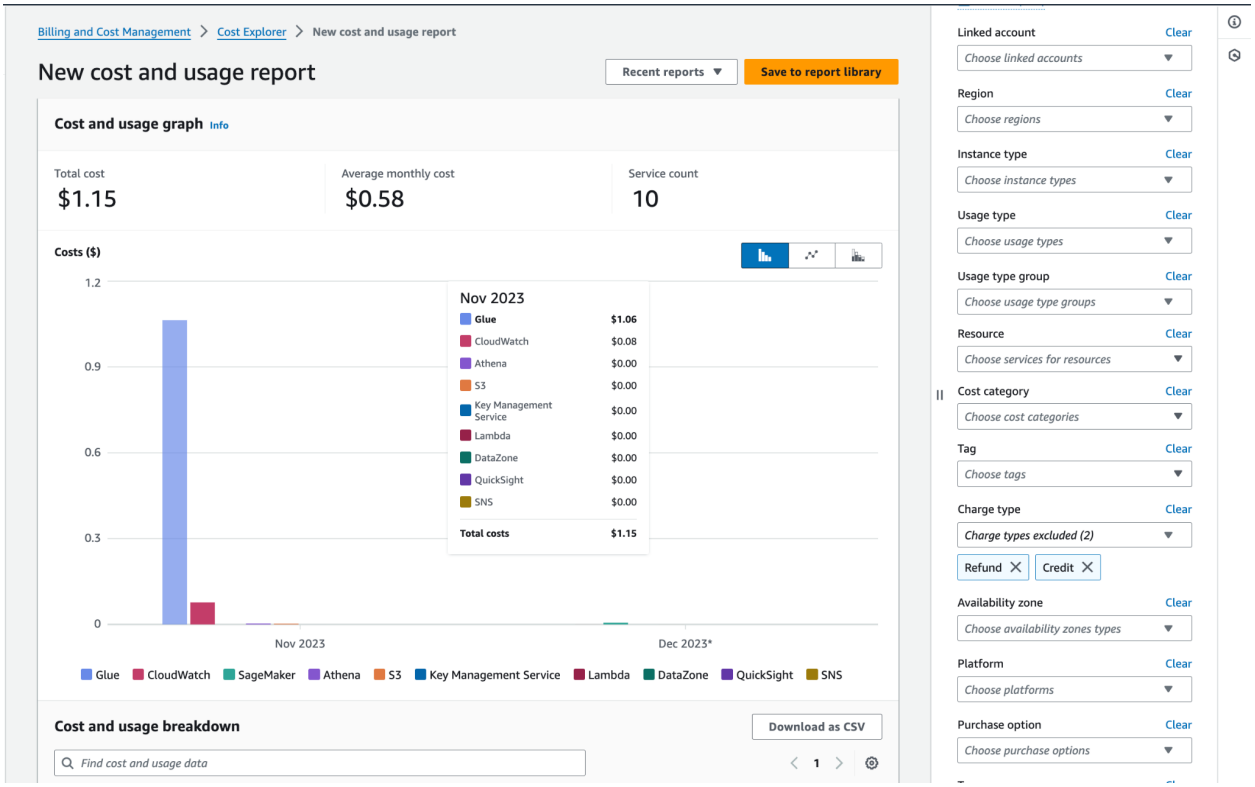
Billing and Cost Management Dashboard



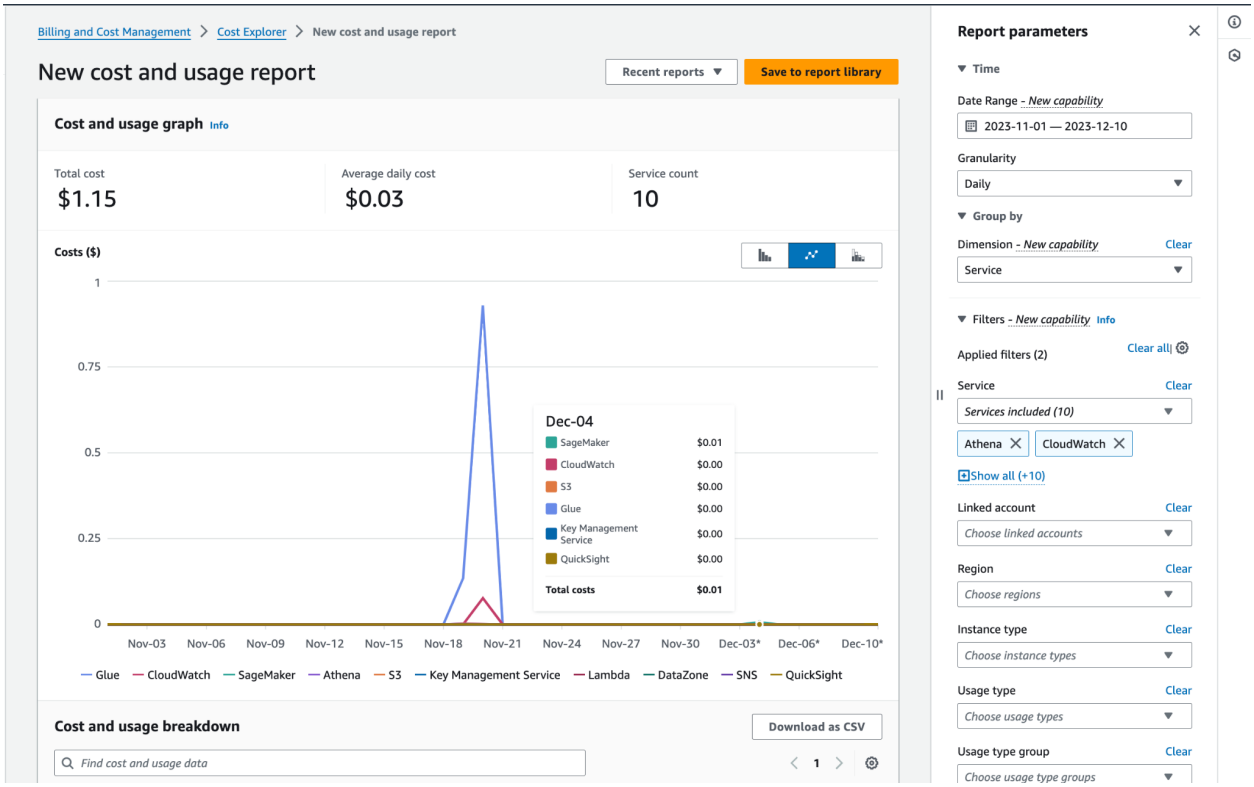
The AWS Billing and Cost Management dashboard presented offers concise yet informative information on current and historical cloud expenditures. The displayed billing and cost management dashboard shows that we have accrued a total of $1.15 in charges over a two-month period. During this timeframe, we have utilized a suite of AWS services, including S3, Glue, Athena, SageMaker, and IAM, among others. This overview encapsulates our financial engagement with AWS's cloud offerings.

New Cost and usage report

The AWS Cost Explorer's 'New Cost and Usage Report' showcases a comprehensive view of **AWS monthly spending**, indicating a total expenditure of $1.15, with AWS Glue service accounting for the majority at $1.06 in November 2023. The average monthly cost is calculated to be $0.58, suggesting variability in monthly expenses. In December 2023, there's a nominal charge of about $0.01 associated with AWS SageMaker, indicating recent or minimal usage of this particular service. Across the board, 10 AWS services are listed, including S3, Athena, QuickSight, CloudWatch, and others. This financial breakdown reflects a concentrated investment in Glue, with a recent engagement in SageMaker, and highlights the potential for optimizing costs in the utilized services. Additionally, the dashboard's functionality to save and generate customized reports allows for effective tracking and management of cloud costs.



The AWS Cost Explorer dashboard vividly illustrates our cloud spend dynamics, reporting a total expenditure of $1.15, averaging a **daily cost** of $0.03 over the period from the start of November to December 10, 2023. Notably, on December 4, there's a recorded expense for AWS SageMaker, albeit a modest $0.01, indicating recent or light usage of the machine learning service. The dashboard encapsulates the activity of 10 AWS services, with a striking spike in costs in November primarily due to AWS Glue. This visualization and the accompanying tools—such as the ability to save reports and the implementation of advanced filters—serve as a powerful aid in tracking and optimizing our cloud spend.



Suggest cost optimization strategies (e.g., using AWS Cost Explorer and AWS Trusted Advisor).

In our comprehensive AWS cost optimization strategy, we meticulously selected resource configurations across various services, including S3, Glue, and SageMaker, to ensure a harmonious balance between performance, resource efficiency, and cost savings.

For our storage needs, we opted for a General Purpose S3 bucket type. This choice was pivotal in optimizing our memory requirements. The General Purpose S3 bucket provided us with a versatile and cost-effective solution for our storage needs, perfectly suited for our data profile. Our total storage requirement amounted to 98.4MB, distributed across 228 objects. This resulted in an average object size of approximately 441.7KB, a size that the General Purpose S3 bucket handles efficiently, both in terms of performance and cost.

In the area of data integration, we utilized AWS Glue's serverless capabilities for our ETL processes. This choice allowed us to benefit from AWS Glue's scalability and flexibility, which were instrumental in managing our data workflows effectively while also keeping an eye on cost control.

Moreover, in our machine learning endeavors with AWS SageMaker, we specifically chose the notebook instance type ml.t3.medium with a 5GB EBS volume. This setup was a critical part of our strategy, striking a balance between the necessary computational power for our machine learning models and cost-effectiveness. The ml.t3.medium instance type, known for its efficient performance-to-cost ratio, was particularly suitable for our medium-scale machine learning tasks. The 5GB EBS volume complemented this by providing adequate storage for our datasets and models without the additional expenses associated with larger volumes or more powerful instances.