

Data Engineering 101:
Synapse vs Bigquery vs Redshift

Azure Synapse Analytics

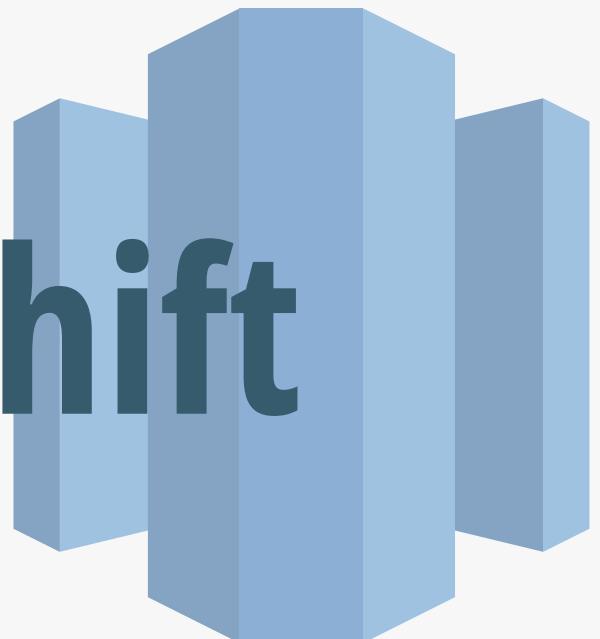
VS



Google BigQuery

VS

Amazon RedShift



Shwetank Singh
GritSetGrow - GSGLearn.com

Data Storage



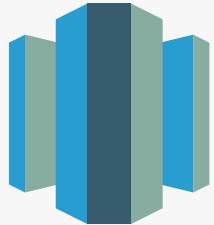
Synapse

Supports SQL Data Warehouse, Data Lake



BigQuery

Columnar storage using Capacitor format



RedShift

Columnar storage with compression techniques



Shwetank Singh
GritSetGrow - GSGLearn.com



Query Language



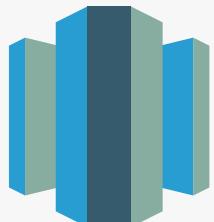
Synapse

T-SQL, Spark SQL, Kusto Query Language



BigQuery

Standard SQL (with extensions)



RedShift

SQL (PostgreSQL-based dialect)



Shwetank Singh
GritSetGrow - GSGLearn.com



Integration



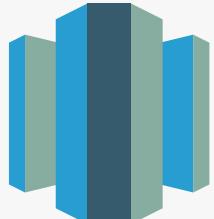
Synapse

Deep integration with Azure ecosystem



BigQuery

Tight integration with Google Cloud services



RedShift

Seamless integration with AWS ecosystem



Shwetank Singh
GritSetGrow - GSGLearn.com



Compute and Storage



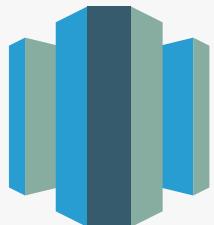
Synapse

Separate compute and storage, serverless options



BigQuery

Serverless, on-demand pricing, scalable storage



RedShift

Compute and storage are separate, with managed options



Shwetank Singh
GritSetGrow - GSGLearn.com



Performance Tuning



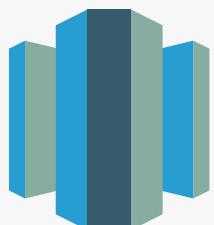
Synapse

Offers automatic and manual tuning options



BigQuery

Automatic performance tuning with no indexes required



RedShift

Requires manual tuning, including sort and distribution keys



Shwetank Singh
GritSetGrow - GSGLearn.com



Pricing Model



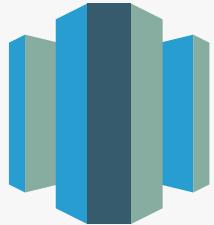
Synapse

Pay-as-you-go, provisioned capacity, serverless



BigQuery

Pay-as-you-go, flat-rate pricing



RedShift

On-demand or reserved instances, predictable
pricing



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Lake Integration



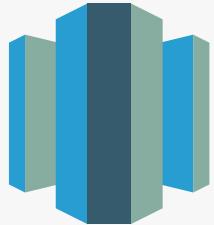
Synapse

Tight integration with Azure Data Lake Storage



BigQuery

Can directly query data in Google Cloud Storage



RedShift

Can query data in Amazon S3 with Spectrum



Shwetank Singh
GritSetGrow - GSGLearn.com



Security



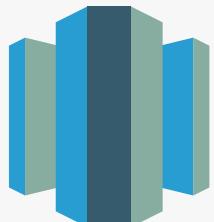
Synapse

Azure AD, RBAC, encryption at rest and in transit



BigQuery

Google IAM, encryption by default



RedShift

IAM, VPC, encryption at rest and in transit



Shwetank Singh
GritSetGrow - GSGLearn.com



Machine Learning Integration



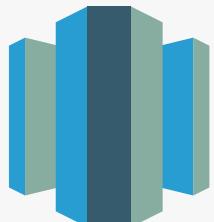
Synapse

Integrated with Azure Machine Learning, Synapse Studio



BigQuery

Integrated with Google AI Platform



RedShift

Integrated with Amazon SageMaker



Shwetank Singh
GritSetGrow - GSGLearn.com



Concurrency



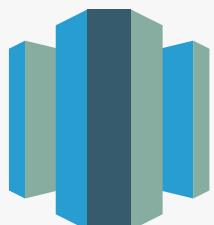
Synapse

Supports high concurrency with Synapse SQL pools



BigQuery

Highly scalable with automatic concurrency management



RedShift

Concurrency scaling via dedicated clusters



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Partitioning



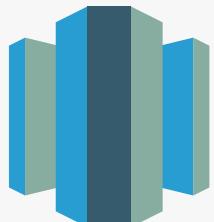
Synapse

Automatic and manual partitioning, including sharding



BigQuery

Automatic partitioning using partitioned tables



RedShift

Requires manual distribution and partitioning



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Ingestion



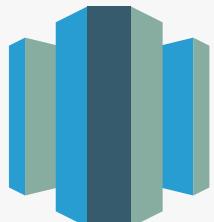
Synapse

Supports batch, real-time, and streaming ingestion



BigQuery

Supports batch and streaming ingestion



RedShift

Supports batch and streaming ingestion



Shwetank Singh
GritSetGrow - GSGLearn.com



Analytics Engine



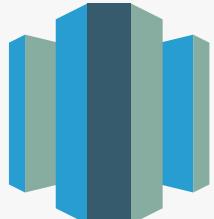
Synapse

Combines SQL, Spark, and Data Explorer engines



BigQuery

Optimized for analytical queries, serverless engine



RedShift

OLAP engine optimized for complex queries



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Processing



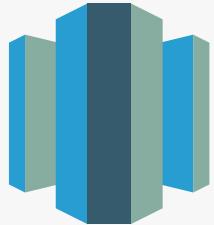
Synapse

Supports batch, streaming, and real-time data processing



BigQuery

Optimized for real-time and batch processing



RedShift

Focuses on batch processing



Shwetank Singh
GritSetGrow - GSGLearn.com



Cross-Region Replication



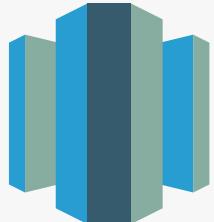
Synapse

Geo-redundant storage, cross-region failover



BigQuery

Multi-region replication, high availability



RedShift

Cross-region snapshot copy



Shwetank Singh
GritSetGrow - GSGLearn.com



Backup and Restore



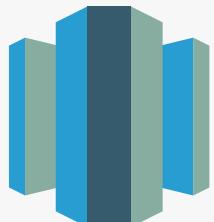
Synapse

Automated backups, point-in-time restore



BigQuery

Automated backups, point-in-time recovery



RedShift

Automated backups, manual snapshots



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Sharing



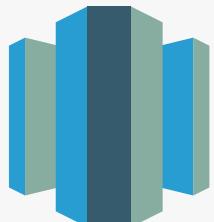
Synapse

Supports data sharing across Azure subscriptions



BigQuery

Data sharing within and across Google Cloud projects



RedShift

Redshift Data Sharing, AWS Data Exchange



Shwetank Singh
GritSetGrow - GSGLearn.com



Scaling



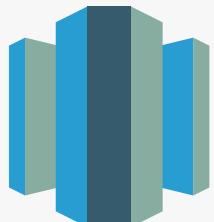
Synapse

Scale compute independently of storage



BigQuery

Scales automatically based on query load



RedShift

Elastic resize, concurrency scaling



Shwetank Singh
GritSetGrow - GSGLearn.com



ETL/ELT Integration



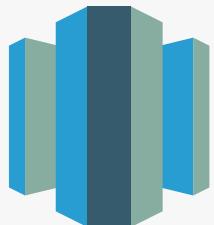
Synapse

Integrated with Azure Data Factory and Data Bricks



BigQuery

Integrated with Google Dataflow and Data Fusion



RedShift

Integrated with AWS Glue and Data Pipeline



Shwetank Singh
GritSetGrow - GSGLearn.com



Cost Optimization



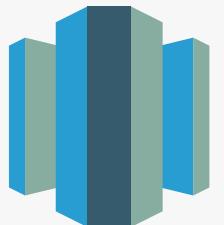
Synapse

Offers reserved instances and serverless options



BigQuery

Supports flat-rate and on-demand pricing models



RedShift

Reserved instances and on-demand pricing



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Engineering 101:
Synapse vs Bigquery vs Redshift

Data Governance



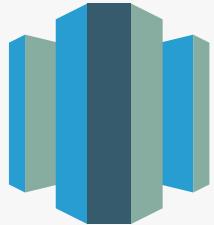
Synapse

Integrated with Azure Purview



BigQuery

Supports Google Cloud Data Catalog



RedShift

Integrated with AWS Lake Formation and IAM



Shwetank Singh
GritSetGrow - GSGLearn.com



Query Performance



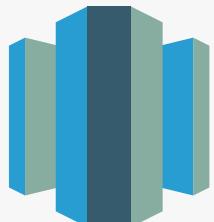
Synapse

Offers materialized views, result caching



BigQuery

Supports materialized views, result caching



RedShift

Supports materialized views, result caching



Shwetank Singh
GritSetGrow - GSGLearn.com



BI Tool Integration



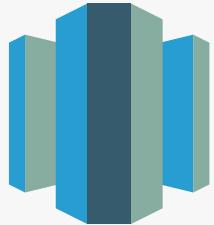
Synapse

Integrated with Power BI



BigQuery

Integrated with Google Data Studio and Looker



RedShift

Integrated with Amazon QuickSight and third-party
BI tools



Shwetank Singh
GritSetGrow - GSGLearn.com



Real-time Analytics



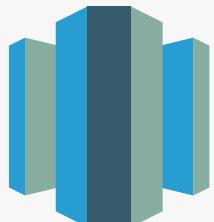
Synapse

Supports real-time analytics with Azure Stream Analytics



BigQuery

Real-time processing with BigQuery



RedShift

Supports real-time analytics with Kinesis and Redshift



Shwetank Singh
GritSetGrow - GSGLearn.com



Multicloud Support



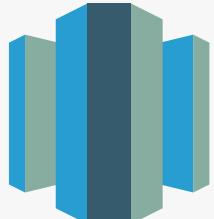
Synapse

Can connect to other clouds via Azure Arc



BigQuery

Supports multicloud via Anthos



RedShift

Primarily focused on AWS, limited multicloud options



Shwetank Singh
GritSetGrow - GSGLearn.com



Compliance and Certifications



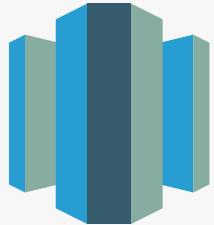
Synapse

Complies with major standards like GDPR, HIPAA



BigQuery

Complies with major standards like GDPR, HIPAA



RedShift

Complies with major standards like GDPR, HIPAA



Shwetank Singh
GritSetGrow - GSGLearn.com



Ecosystem



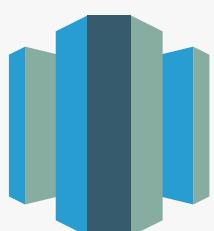
Synapse

Part of Azure Synapse workspace, integrates multiple services



BigQuery

Part of Google Cloud Platform ecosystem, integrates multiple services



RedShift

Part of AWS ecosystem, integrates multiple services



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Warehousing



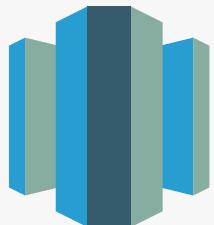
Synapse

Enterprise-scale data warehousing with complex queries



BigQuery

Large-scale data warehousing with serverless flexibility



RedShift

OLAP workloads with complex queries and large datasets



Shwetank Singh
GritSetGrow - GSGLearn.com



Real-time Analytics



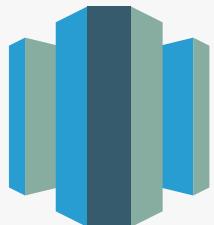
Synapse

Real-time analytics and dashboards in Synapse Studio



BigQuery

Real-time analytics on streaming data



RedShift

Real-time data processing and analytics using Kinesis



Shwetank Singh
GritSetGrow - GSGLearn.com



Machine Learning



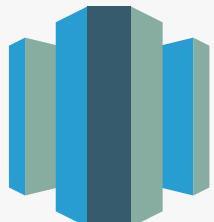
Synapse

Advanced analytics and machine learning with
Azure ML



BigQuery

Machine learning at scale with BigQuery ML



RedShift

Machine learning with Amazon SageMaker



Shwetank Singh
GritSetGrow - GSGLearn.com



Big Data Analytics



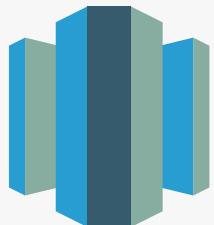
Synapse

Big data processing and analysis using Spark and SQL



BigQuery

Scalable big data analytics with built-in ML features



RedShift

Processing and analyzing large datasets with OLAP capabilities



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Integration



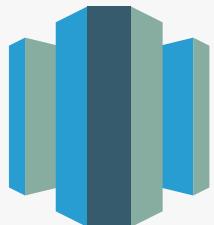
Synapse

Integration with various data sources using Data Factory



BigQuery

Integration with multiple Google Cloud services



RedShift

Integration with AWS data sources using AWS Glue



Shwetank Singh
GritSetGrow - GSGLearn.com



Reporting and Visualization



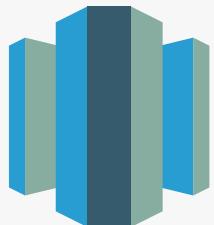
Synapse

Creating reports and dashboards using Power BI



BigQuery

Integrated with Google Data Studio for reporting



RedShift

BI reporting with Amazon QuickSight and third-party tools



Shwetank Singh
GritSetGrow - GSGLearn.com



Cross-Cloud Data Sharing



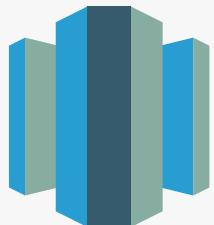
Synapse

Sharing data across different Azure regions and clouds



BigQuery

Seamless data sharing across Google Cloud projects



RedShift

Sharing data across AWS regions and accounts



Shwetank Singh
GritSetGrow - GSGLearn.com



Disaster Recovery



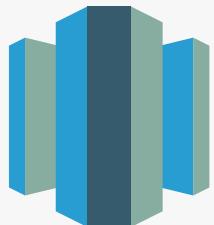
Synapse

Cross-region replication and geo-redundancy for DR



BigQuery

Multi-region availability and replication for DR



RedShift

Cross-region replication for disaster recovery



Shwetank Singh
GritSetGrow - GSGLearn.com



Data Governance



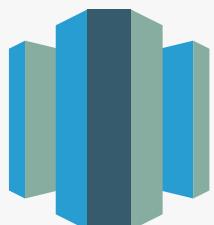
Synapse

Managing and governing data with Azure Purview



BigQuery

Data governance and cataloging with Google Data Catalog



RedShift

Governance and cataloging with AWS Lake Formation



Shwetank Singh
GritSetGrow - GSGLearn.com

