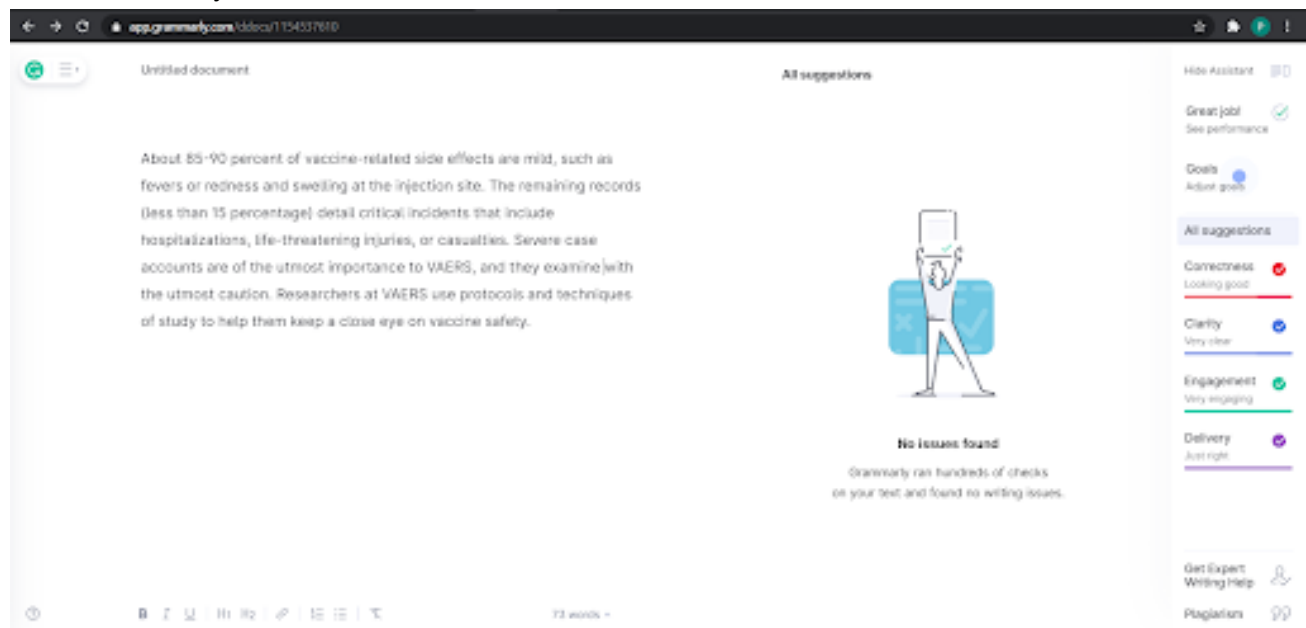


Used Grammarly:



The screenshot shows the Grammarly web application interface. At the top, the browser address bar displays 'app.grammarly.com/doc/115437610'. The main editing area contains the following text: 'About 85-90 percent of vaccine-related side effects are mild, such as fevers or redness and swelling at the injection site. The remaining records (less than 15 percentage) detail critical incidents that include hospitalizations, life-threatening injuries, or casualties. Severe case accounts are of the utmost importance to VAERS, and they examine with the utmost caution. Researchers at VAERS use protocols and techniques of study to help them keep a close eye on vaccine safety.' Below the text, a large blue checkmark icon is displayed with the heading 'No issues found' and the message 'Grammarly ran hundreds of checks on your text and found no writing issues.' On the right side, a sidebar titled 'All suggestions' shows a 'Great job!' message and a list of metrics: 'Correctness' (Looking good), 'Clarity' (Very clear), 'Engagement' (Very engaging), and 'Delivery' (Just right). At the bottom of the interface, a word count of '73 words' is visible.

An entity-relationship diagram depicts the relationships between entity sets contained in a database (ERD). In this way, an individual, or a data unit, is an entity. The term "entity package" refers to a collection of similar entities. Attributes are used to specify properties for these entities.



This screenshot shows a second instance of the Grammarly web application. The main editing area is currently blank. A large blue checkmark icon is displayed with the heading 'No issues found' and the message 'Grammarly ran hundreds of checks on your text and found no writing issues. Check back in when you're ready to write some more.' The bottom of the interface shows a word count of '45 words'.

AWS implementation:

S3 Bucket:

The screenshot shows the Amazon S3 console interface. On the left is a navigation sidebar with 'Amazon S3' selected. The main area displays the 'covidadversebucket' with tabs for 'Objects', 'Properties', 'Permissions', 'Metrics', 'Management', and 'Access Points'. The 'Objects' tab is active, showing a list of 5 objects. Above the list are buttons for 'Copy URL', 'Open', 'Download', 'Delete', 'Actions', 'Create folder', and 'Upload'. A search bar is present above the table. The table lists the following objects:

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	patient_history_org.csv	csv	May 6, 2021, 01:21:24 (UTC-07:00)	5.8 MB	Standard
<input type="checkbox"/>	Patient_Info.csv	csv	May 6, 2021, 01:21:26 (UTC-07:00)	1.8 MB	Standard
<input type="checkbox"/>	patient_outcomes.csv	csv	May 6, 2021, 01:21:25 (UTC-07:00)	1.9 MB	Standard
<input type="checkbox"/>	symptoms.csv	csv	May 6, 2021, 01:21:29 (UTC-07:00)	4.9 MB	Standard
<input type="checkbox"/>	vaccine_details.csv	csv	May 6, 2021, 01:21:30 (UTC-07:00)	1.9 MB	Standard

AWS Glue:

Jobs for each table:

AWS Glue

Search for services, features, marketplace products, and docs [Option+S]

PravallikaV Oregon Support

AWS Glue

Data catalog

Databases

Tables

Connections

Crawlers

Classifiers

Schema registries

Schemas

Settings

ETL

AWS Glue Studio **new**

Blueprints

Workflows

Jobs

ML Transforms

Triggers

Dev endpoints

Notebooks

Security

Security configurations

Tutorials

Add crawler

Explore table

Tables A table is the metadata definition that represents your data, including its schema. A table can be used as a source or target in a job definition.

Add tables Action Database: csvdatafroms3 Filter or search for tables... Save view Showing: 1 - 5

Name	Database	Location	Classification	Last updated	Deprecated
vaccine_details_csv	csvdatafroms3	s3://covidadversebucket/vaccin...	csv	6 May 2021 3:47 AM UTC-7	
patient_outcomes_csv	csvdatafroms3	s3://covidadversebucket/patient...	csv	6 May 2021 3:47 AM UTC-7	
symptoms_csv	csvdatafroms3	s3://covidadversebucket/sympto...	csv	6 May 2021 3:47 AM UTC-7	
patient_history_org_csv	csvdatafroms3	s3://covidadversebucket/patient...	csv	6 May 2021 3:47 AM UTC-7	
patient_info_csv	csvdatafroms3	s3://covidadversebucket/Patient...	csv	5 May 2021 6:50 PM UTC-7	

© 2008 - 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

Glue Connections:

RedshiftCluster: JDBC connection to send the tables into redshift db

S3connection: Network connection to fetch the raw data from s3

AWS Glue

Search for services, features, marketplace products, and docs [Option+S]

PravallikaV Oregon Support

AWS Glue

Data catalog

Databases

Tables

Connections

Crawlers

Classifiers

Schema registries

Schemas

Settings

ETL

AWS Glue Studio **new**

Blueprints

Workflows

Jobs

ML Transforms

Triggers

Dev endpoints

Notebooks

Security

Security configurations

Tutorials

Add crawler

Explore table

Connections A connection contains the properties needed to connect to your data.

Add connection Test connection Action Showing: 1 - 2

Name	Type	Date created	Last updated	Updated by
RedShiftCluster	JDBC	6 May 2021 10:48 AM UTC-7	6 May 2021 3:15 PM UTC-7	root
s3connection	Network	6 May 2021 2:59 PM UTC-7	6 May 2021 2:59 PM UTC-7	root

https://us-west-2.console.aws.amazon.com/glue/home?region=us-west-2#catalog:tab=connections

© 2008 - 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

Crawler:

The screenshot shows the AWS Glue console interface. On the left is a navigation menu with categories like Data catalog, Databases, ETL, and Security. The main content area is titled 'Crawlers' and includes a description: 'A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.' Below this is a table listing crawlers. One crawler, 'covidddb225', is shown with a status of 'Ready', a last runtime of '1 min', and 5 tables added.

Name	Schedule	Status	Logs	Last runtime	Median runtime	Tables updated	Tables added
covidddb225		Ready	Logs	1 min	1 min	0	5

VPC Endpoint:

The screenshot shows the AWS VPC console interface. The main content area is titled 'Resources by Region' and displays a grid of VPC resources for the 'Oregon' region. Resources include VPCs (1), Subnets (4), Route Tables (1), Internet Gateways (1), Egress-only Internet Gateways (0), DHCP options sets (1), Elastic IPs (0), Endpoints (0), NAT Gateways (0), VPC Peering Connections (0), Network ACLs (1), Security Groups (1), Customer Gateways (0), Virtual Private Gateways (0), Site-to-Site VPN Connections (0), and Running Instances (0). The right sidebar contains sections for 'Current status' (showing 'Amazon EC2 - US West (Oregon)' as 'Service is operating normally'), 'Settings', 'Additional information', 'Transit Gateway Network Manager', and 'Site-to-Site VPN Connections'.

Amazon Redshift Cluster:

aws

Services

Search for services, features, marketplace products, and docs

[Option+S]

PravallikaV

Oregon

Support

Amazon Redshift

Clusters

redshift-covid-cluster

redshift-covid-cluster

Actions

Edit

Add partner integration

Query cluster

General information

Cluster identifier

redshift-covid-cluster

Cluster namespace

53f24b88-65eb-46bd-bd81-6aa6d0f28574

Status

Paused

Date created

May 06, 2021, 12:00(UTC-07:00)

Storage used

-

Node type

dc2.large

Number of nodes

1

AQUA

Not available

Endpoint

redshift-covid-cluster.cbpgvcniwu5f.us-west-2.r...

JDBC URL

jdbc:redshift://redshift-covid-cluster.cbpgvcniw...

ODBC URL

Driver={Amazon Redshift (x64)}; Server=redshif...

Cluster performance

Query monitoring

Schedules

Maintenance

Properties

Database configurations

Change admin user password

Rotate encryption keys

Edit

Database name

dev

Parameter group

Defines database parameter and query queues for all the databases.
default.redshift-1.0

Encryption

Disabled

Audit logging

Disabled

Port

5439

AWS KMS key ID

-

SSH ingestion setting (cluster public key)

Feedback

English (US)

© 2008 - 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Cookie preferences

DynamoDB table:

aws

Services

Search for services, features, marketplace products, and docs

[Option+S]

PravallikaV

Oregon

Support

DynamoDB

Dashboard

Tables

Backups

Reserved capacity

Exports to S3

Item explorer

PartiQL editor

DAX

Revert to the current console

DynamoDB

Tables

Tables (1)

Info

Find tables by table name

Any group

1

Name

Status

Partition key

Sort key

Indexes

Read capacity mode

Write capacity mode

CovidTable

Active

patient_no (String)

-

0

Provisioned with auto scaling (5)

Provisioned with auto scaling (5)

Feedback

English (US)

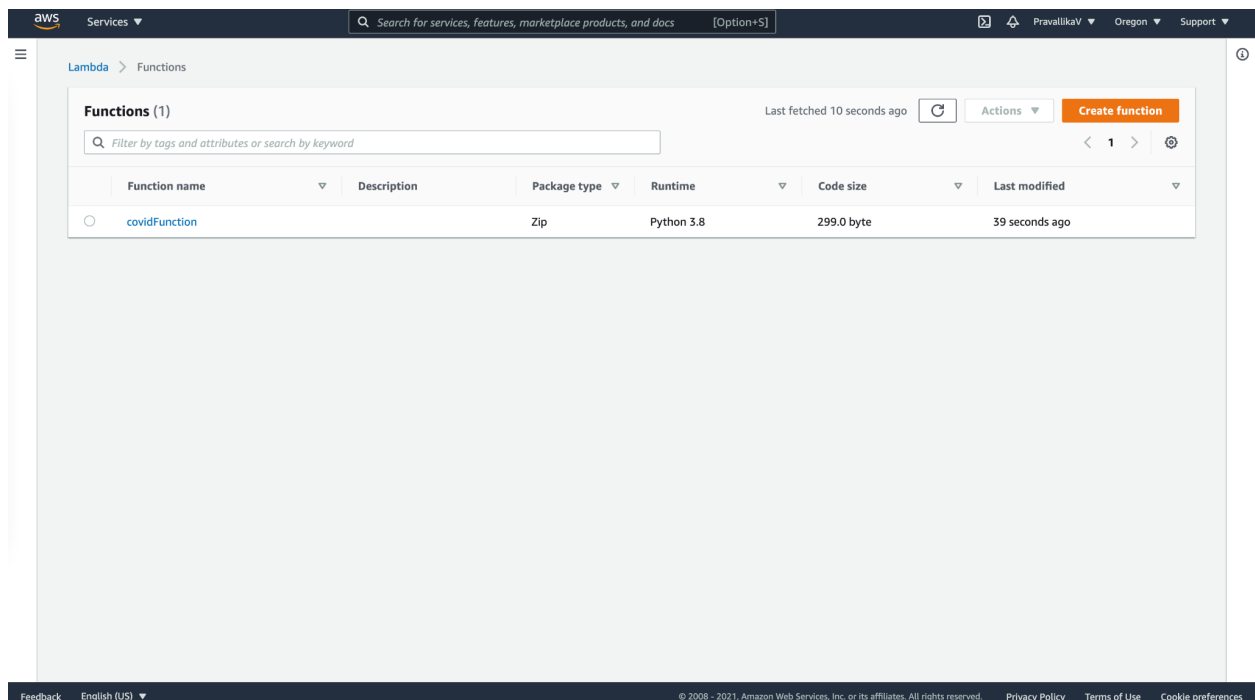
© 2008 - 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Cookie preferences

Lambda Function:



Used Scrum methodology to

A	B	C
Date	Link	MOM
March 5	https://meet.google.com/tcn-cgkr-wmg?hs=151	Discussed about the project prosal about requirements and tools to implement
March 19	https://meet.google.com/htz-vkft-hdk	Modeled Data and created tables accordingly for the project using workbench and postgre
April 3	https://sjsu.zoom.us/j/85858930125?pwd=c2VN	Discussed about the data cleaning and tools to be used for data cleaning
April 11	https://meet.google.com/bmx-encj-jsn	Discussion for intermediate Status report and Further deliverables to be implemnted
April 14	in-person on campus	Discussion about the project flow
April 20	https://meet.google.com/oed-xdww-seg	Made python connectivity with postgre and performed some SQL queries and started working to move the data to AWS
April 28	in-person on campus	Discussion about overall progress of project
April 29	https://meet.google.com/hjf-pdlip-ekf	moved data s3 and performed some glue jobs to move to redshift and similary did in dynamdb
May 1	https://us04web.zoom.us/j/7070941253?pwd=W	Discussion about significant case study and procedure to implement
May 3	on campus	Discussion about industry case study and procedure to implement
May 5	on campus	Perfomed analysis using tableau and seaborn library and started working on both case study and significant paper
May 7	https://meet.google.com/czp-pwnk-pld?hs=151	Started our Project report and presentation