

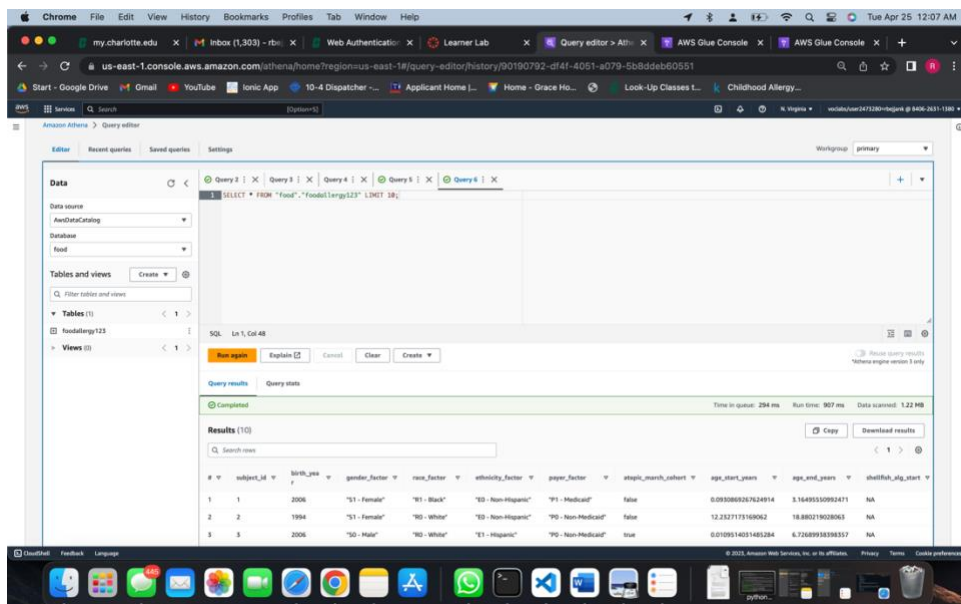
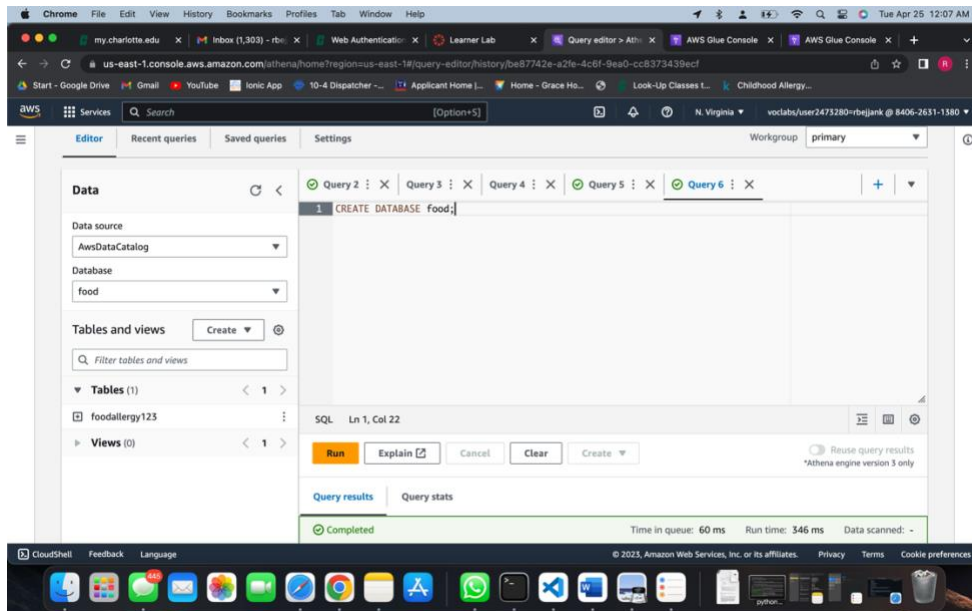
Data understanding:

a. Exploratory Data Analysis

AWS components used:

Amazon S3 bucket - s3://foodallergy123/foodAllergy.csv

Amazon Athena:



Chrome File Edit View History Bookmarks Profiles Tab Window Help

my.charlotte.edu x Inbox (1,303) - rbe... x Web Authentication: x Learner Lab x Query editor > Ath... x AWS Glue Console x AWS Glue Console x +

us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/da86098f-501e-42d5-b757-814cb34809e9

Start - Google Drive Gmail YouTube Ionic App 10-4 Dispatcher ... Applicant Home ... Home - Grace Ho... Look-Up Classes t... Childhood Allergy...

AWS Services Search [Options+5]

Amazon Athena > Query editor

Editor Recent queries Saved queries Settings Workgroup primary

Data

Data source: AwsDataCatalog Database: food

Tables and views

Tables (1) foodallergy123 Views (0)

SQL Ln 1, Col 94

```
1 select birth_year,max(gender_factor) as gender_value from foodallergy123 group by birth_year;
```

Run again Explain Cancel Clear Create

Query results Query stats

Completed Time in queue: 141 ms Run time: 1.265 sec Data scanned: 83.68 MB

Results (30)

Search rows

#	birth_year	gender_value
1	2000	"S1 - Female"
2	1989	"S1 - Female"
3	2008	"S1 - Female"
4	1998	"S1 - Female"

© 2023 Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Chrome File Edit View History Bookmarks Profiles Tab Window Help

my.charlotte.edu x Inbox (1,303) - rbe... x Web Authentication: x Learner Lab x Query editor > Ath... x AWS Glue Console x AWS Glue Console x +

us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/826205cd-c024-48b4-9c48-353010da0347

Start - Google Drive Gmail YouTube Ionic App 10-4 Dispatcher ... Applicant Home ... Home - Grace Ho... Look-Up Classes t... Childhood Allergy...

AWS Services Search [Options+5]

Amazon Athena > Query editor

Editor Recent queries Saved queries Settings Workgroup primary

Data

Data source: AwsDataCatalog Database: food

Tables and views

Tables (1) foodallergy123 Views (0)

SQL Ln 1, Col 61

```
1 select sum (age_start_years) as age_sum from foodallergy123;
```

Run again Explain Cancel Clear Create

Query results Query stats

Completed Time in queue: 179 ms Run time: 976 ms Data scanned: 83.68 MB

Results (1)

Search rows

#	age_sum
1	1313521.182751378

© 2023 Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Chrome File Edit View History Bookmarks Profiles Tab Window Help

my.charlotte.edu x Inbox (1,303) - rbe... x Web Authentication: x Learner Lab x Query editor > Ath... x AWS Glue Console x AWS Glue Console x +

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/crawlers/

Start - Google Drive Gmail YouTube Ionic App 10-4 Dispatcher ... Applicant Home ... Home - Grace Ho... Look-Up Classes t... Childhood Allergy...

Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)

▼ Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings

▼ Data Integration and ETL
AWS Glue Studio
Jobs
Interactive Sessions
Notebooks
Data classification tools
Sensitive data detection
Record Matching
Triggers
Workflows (orchestration)
Blueprints
Security configurations

► Legacy pages

What's New
Documentation
AWS Marketplace

Enable compact mode
Enable new navigation

CloudShell Feedback Language

Crawler successfully starting
The following crawler is now starting: "foodallergy123"

AWS Glue > Crawlers

Crawlers
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.

Crawlers (1) Info
View and manage all available crawlers.
Last updated: 4/25/2023 at 04:14:03

Filter crawlers

Name	State	Schedule	Last run	Last run time	Log	Table changes from last run
foodallergy123	Ready		Succeeded	April 25, 2023 at 4:00:05	View log	1 created

Chrome File Edit View History Bookmarks Profiles Tab Window Help

my.charlotte.edu x Inbox (1,303) - rbe... x Web Authentication: x Learner Lab x Query editor > Ath... x AWS Glue Console x AWS Glue Console x +

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/tables/view/foodallergy123?database=food&catalogId=840626311380&versionId=...

Start - Google Drive Gmail YouTube Ionic App 10-4 Dispatcher ... Applicant Home ... Home - Grace Ho... Look-Up Classes t... Childhood Allergy...

Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)

▼ Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings

▼ Data Integration and ETL
AWS Glue Studio
Jobs
Interactive Sessions
Notebooks
Data classification tools
Sensitive data detection
Record Matching
Triggers
Workflows (orchestration)
Blueprints
Security configurations

► Legacy pages

What's New
Documentation
AWS Marketplace

Enable compact mode
Enable new navigation

CloudShell Feedback Language

Name: foodallergy123
Description: -
Database: food
Classification: csv
Location: s3://foodallergy123/
Connection: -
Input format: org.apache.hadoop.mapred.TextInputFormat
Output format: org.apache.hadoop.hive ql.io.HiveIgnoreKeyTextOutputFormat
Send serialization lib: org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe

Schema Partitions Indexes

Schema (50)
View and manage the table schema.
Edit schema as JSON Edit schema

Filter schemas

#	Column name	Data type	Partition key	Comment
1	subject_id	bigint	-	-
2	birth_year	bigint	-	-
3	gender_factor	string	-	-
4	race_factor	string	-	-
5	ethnicity_factor	string	-	-
6	payer_factor	string	-	-
7	atopic_march_cohort	boolean	-	-
8	age_start_years	double	-	-
9	age_end_years	double	-	-
10	shellfish_allg_start	string	-	-
11	shellfish_allg_end	string	-	-
12	fish_allg_start	string	-	-
13	fish_allg_end	string	-	-
14	milk_allg_start	string	-	-
15	milk_allg_end	string	-	-
16	soy_allg_start	string	-	-