**create** **table** sourc\_dim\_dist(

Source\_IP **varchar**(200),

Source\_Port **varchar**(200),

surrogate\_key **int** **primary** **key**,

Source\_IP\_Port **varchar**(400)

) **distkey**(Source\_IP)

**sortkey** (Source\_Port);

**copy** sourc\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Source\_IP.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** sourc\_dim\_dist **limit** 5;

**create** **table** dest\_dim\_dist(

Destination\_IP **varchar**(200),

Destination\_Port **varchar**(200),

surrogate\_key **int** **primary** **key**,

Destination\_IP\_Port **varchar**(400)

)**distkey**(Destination\_IP)

**sortkey**(Destination\_Port);

**copy** dest\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Destination\_IP.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** dest\_dim\_dist **limit** 5;

**create** **table** protocol\_dim\_dist(

Protocol **int**,

Protocol\_Name **varchar**(10),

Protocol\_Description **varchar**(300),

surrogate\_key **int** **primary** **key**

)**distkey**(Protocol);

**copy** protocol\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Protocol.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** protocol\_dim\_dist;

**create** **table** Packet\_dim\_dist(

Total\_Fwd\_Packets **float**,

Total\_Bwd\_Packets **float**,

Total\_Length\_Fwd\_Packets **float**,

Total\_Length\_Bwd\_Packets **float**,

Fwd\_Packet\_Length\_Max **float**,

Fwd\_Packet\_Length\_Min **float**,

Fwd\_Packet\_Length\_Mean **float**,

Fwd\_Packet\_Length\_Std **float**,

Bwd\_Packet\_Length\_Max **float**,

Bwd\_Packet\_Length\_Min **float**,

Bwd\_Packet\_Length\_Mean **float**,

Bwd\_Packet\_Length\_Std **float**,

Min\_packet\_Length **float**,

Max\_packet\_Length **float**,

Packet\_Length\_Mean **float**,

Packet\_Length\_Std **float**,

Packet\_Length\_Var **float**,

Average\_Packet\_Size **float**,

Combined\_Col **varchar**(500),

surrogate\_key **int** **primary** **key**

)**distkey**(surrogate\_key);

**copy** Packet\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Packet.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** Packet\_dim\_dist **limit** 1;

**create** **table** flag\_dim\_dist(

Fwd\_Psh\_flags **int**,

Bwd\_Psh\_flags **int**,

Fwd\_Urg\_Flags **int**,

Bwd\_Urg\_Flags **int**,

FIN\_Flag\_Count **int**,

SYN\_Flag\_Count **int**,

RST\_Flag\_Count **int**,

PSH\_Flag\_Count **int**,

ACK\_Flag\_Count **int**,

URG\_Flag\_Count **int**,

CWE\_Flag\_Count **int**,

ECE\_Flag\_Count **int**,

Combined\_Col **varchar**(500),

surrogate\_key **int** **primary** **key**

)**distkey**(surrogate\_key);

**copy** flag\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Flag.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** flag\_dim\_dist **limit** 2;

**create** **table** date\_dim\_dist(

Time\_Stamp **timestamp**,

**day** **int**,

**month** **int**,

**year** **int**,

**time** **varchar**(200),

surrogate\_key **int** **primary** **key**

)**distkey**(**year**)

**sortkey**(surrogate\_key);

**copy** date\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/Date.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** \* **from** date\_dim\_dist **limit** 2;

**create** **table** flow\_dim\_dist(

Sr\_No **int**,

Flow\_ID **varchar**(500),

Flow\_Duration **int**,

Flow\_Bytes\_per\_sec **float**,

Flow\_Packets\_per\_s **float**,

Flow\_IAT\_Mean **float**,

Fwd\_IAT\_Std **float**,

Fwd\_IAT\_Max **float**,

Fwd\_IAT\_Min **float**,

Merged\_Col **varchar**(300),

surrogate\_key **int** **primary** **key**

)**distkey**(Flow\_ID)

**sortkey**(surrogate\_key);

**copy** flow\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/flow\_1.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** flow\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/flow\_2.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** flow\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/flow\_3.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** flow\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/flow\_4.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** flow\_dim\_dist **from** 's3://myawsbucketfordwprojectwest/flow\_5.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**select** **count**(\*) **from** flow\_dim\_dist;

**create** **table** fact\_table\_dist(

Sr\_No **int**,

Source\_ID\_FK **int**,

Dest\_IP\_FK **int**,

Protocol\_FK **int**,

Timestamp\_FK **int**,

Packet\_FK **int**,

Flow\_FK **int**,

Flag\_FK **int**,

Label **varchar**(200)

)**distkey**(Label)

**sortkey**(Sr\_No);

**copy** fact\_table\_dist **from** 's3://myawsbucketfordwprojectwest/fact\_table\_1.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** fact\_table\_dist **from** 's3://myawsbucketfordwprojectwest/fact\_table\_2.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**copy** fact\_table\_dist **from** 's3://myawsbucketfordwprojectwest/fact\_table\_3.csv' IAM\_ROLE 'arn:aws:iam::897729099108:role/service-role/AmazonRedshift-CommandsAccessRole-20241023T060326'

**csv** **ignoreheader** 1;

**SET** enable\_result\_cache\_for\_session **TO** **OFF**;

--FTP-Patator Attack Frequency by Source Details

**select** s.source\_ip,s.source\_port,f.Label,**count**(\*) **from** sourc\_dim\_dist s **join** fact\_table\_dist f **on** s.surrogate\_key=f.Source\_ID\_FK

**group** **by** (s.source\_ip,s.source\_port,f.Label) **having** f.Label='FTP-Patator';

--Query to Retrieve Top Source IP and Port by Event Count for Each Label

**with** *labelcount* **as** (

**select** s.source\_ip,s.source\_port,f.Label,**count**(\*) **as** cnt,

**row\_number**() **over** (**partition** **by** f.Label **order** **by** cnt **desc** ) **as** row\_rank **from** sourc\_dim\_dist s **join** fact\_table\_dist f

**on** s.surrogate\_key=f.Source\_ID\_FK **group** **by** s.source\_ip,s.source\_port,f.Label

)

**select** source\_ip,source\_port,Label,*cnt* **as** *Max\_Count* **from** *labelcount* **where** *row\_rank*=1;

--Average Packet Details for each attack type

**select** **avg**(p.Average\_Packet\_Size) **as** Average\_Packet\_Size,f.Label **from** Packet\_dim\_dist p **join** fact\_table\_dist f

**on** f.Packet\_FK=p.surrogate\_key **group** **by** (f.Label) **order** **by** Average\_Packet\_Size;

--Hourly Distribution of Traffic for Each Attack Type

**select** f.Label,**extract**(**hour** **from** d.time\_stamp) **as** **Hour**,**count**(\*) **as** Total\_Count **from** fact\_table\_dist f **join** date\_dim\_dist d

**on** f.Timestamp\_FK=d.surrogate\_key **group** **by** f.Label,**hour** **order** **by** Total\_Count **desc**;

--Identify Top 3 Destination IPs with Highest Traffic Count for Each Label

**with** denserankcount **as** (

**select** *d*.destination\_ip,*f*.***Label***,**count**(\*) **as** *Total\_Count*,**dense\_rank**() **over**(**partition** **by** *f*.***Label*** **order** **by** Total\_Count **desc**)

**from** ***dest\_dim\_dist*** **as** *d* **join** fact\_table\_dist **as** *f* **on** *f*.***Dest\_IP\_FK***=*d*.surrogate\_key **group** **by** *d*.destination\_ip,*f*.***Label***)

**select** destination\_ip,Label,Total\_Count **from** ***denserankcount*** **where** dense\_rank<=3;

--Total Packets Forwarded by Source IP and Protocol

**select** *s*.***source\_ip***, *p*.protocol\_name,**sum**(*pa*.Total\_Fwd\_Packets) **as** *Total\_Packet\_Forwarded* **from** sourc\_dim\_dist *s* **join** fact\_table\_dist *f*

**on** *f*.***Source\_ID\_FK***=*s*.***surrogate\_key***

**join** ***Protocol\_dim\_dist*** *p* **on** *p*.surrogate\_key=*f*.***Protocol\_FK*** **join** ***Packet\_dim\_dist*** *pa* **on** *pa*.surrogate\_key=*f*.***Packet\_FK***

**group** **by** *s*.***source\_ip***,*p*.protocol\_name **order** **by** *Total\_Packet\_Forwarded* **DESC**;

--Top 1 Total Packets Forwarded by Source IP and Protocol

**with** totalpacketforwarded **as**(

**select** *s*.***source\_ip***,*p*.protocol\_name,**sum**(*pa*.Total\_Fwd\_Packets) **as** *Total\_Packet\_forwarded* ,

**dense\_rank** () **over** (**partition** **by** *s*.***source\_ip***,*p*.protocol\_name **order** **by** ***Total\_Packet\_forwarded*** **DESC**) **as** *dense\_rank* **from** sourc\_dim\_dist *s* **join** fact\_table\_dist *f* **on**

*s*.***surrogate\_key***=*f*.***source\_ID\_FK*** **join** ***Protocol\_dim\_dist*** *p* **on** *p*.surrogate\_key=*f*.***Protocol\_FK*** **join** ***Packet\_dim\_dist*** *pa* **on** *pa*.surrogate\_key=*f*.***Packet\_FK***

**group** **by** *s*.***source\_ip***,*p*.protocol\_name)

**select** source\_ip,protocol\_name,Total\_Packet\_forwarded **from** ***totalpacketforwarded*** **where** dense\_rank=1 **order** **by** Total\_Packet\_forwarded **DESC**;

--Identify Sources with a Packet Size Spike Compared to Previous Entry having size spike of more than 1000

**with** source\_lag **as**(

**select** *s*.***source\_ip***,*d*.time\_stamp,*pa*.Packet\_Length\_Mean,

**lag**(*pa*.Packet\_Length\_Mean,1) **over**(**partition** **by** *s*.***source\_ip*** **order** **by** *d*.time\_stamp) **as** *previous\_value*

**from** sourc\_dim\_dist *s* **JOIN** fact\_table\_dist *f* **ON** *s*.***surrogate\_key*** = *f*.***Source\_ID\_FK***

**JOIN** ***Packet\_dim\_dist*** *pa* **ON** *pa*.surrogate\_key = *f*.***Packet\_FK***

**JOIN** ***date\_dim\_dist*** *d* **ON** *d*.surrogate\_key = *f*.***Timestamp\_FK***)

**select** source\_ip,time\_stamp,packet\_length\_mean,(packet\_length\_mean-previous\_value) **as** *Size\_difference* **from** ***source\_lag***

**where** size\_difference>1000;

--Identifying Source and Destination Combinations with High Flag Count

**SELECT** *s*.***source\_ip***, *d*.destination\_ip, **SUM**(*fl*.FIN\_Flag\_Count + *fl*.SYN\_Flag\_Count + *fl*.PSH\_Flag\_Count) **AS** *total\_flags*

**FROM** sourc\_dim\_dist *s*

**JOIN** fact\_table\_dist *f* **ON** *s*.***surrogate\_key*** = *f*.***Source\_ID\_FK***

**JOIN** ***dest\_dim\_dist*** *d* **ON** *d*.surrogate\_key = *f*.***Dest\_IP\_FK***

**join** ***flag\_dim\_dist*** *fl* **on** *fl*.surrogate\_key=*f*.***FLAG\_FK***

**GROUP** **BY** *s*.***source\_ip***, *d*.destination\_ip

**HAVING** ***total\_flags*** > 100;

--Anomaly Detection in Network Traffic Using Flag Count and Hourly Analysis

**WITH** *TrafficAnomalies\_dist* **AS** (

**SELECT**

*s*.***source\_ip***,

*d*.destination\_ip,

*p*.protocol\_name,

**EXTRACT**(**hour** **FROM** *t*.time\_stamp) **AS** *hour*,

**SUM**(*fl*.SYN\_Flag\_Count + *fl*.ACK\_Flag\_Count) **AS** *total\_flags*,

**AVG**(*pa*.Packet\_Length\_Mean) **AS** *avg\_packet\_length*,

**COUNT**(\*) **AS** *event\_count*

**FROM**

sourc\_dim\_dist **AS** *s*

**JOIN**

fact\_table\_dist **AS** *f* **ON** *s*.***surrogate\_key*** = *f*.***Source\_ID\_FK***

**JOIN**

***dest\_dim\_dist*** **AS** *d* **ON** *d*.surrogate\_key = *f*.***Dest\_IP\_FK***

**JOIN**

***protocol\_dim\_dist*** **AS** *p* **ON** *p*.surrogate\_key = *f*.***Protocol\_FK***

**JOIN**

***flag\_dim\_dist*** **AS** *fl* **ON** *fl*.surrogate\_key = *f*.***Flag\_FK***

**JOIN**

***packet\_dim\_dist*** **AS** *pa* **ON** *pa*.surrogate\_key = *f*.***Packet\_FK***

**JOIN**

***date\_dim\_dist*** **AS** *t* **ON** *t*.surrogate\_key = *f*.***Timestamp\_FK***

**WHERE**

*p*.protocol\_name **IN** ('TCP', 'UDP')

**GROUP** **BY**

*s*.***source\_ip***, *d*.destination\_ip, *p*.protocol\_name, *hour*

**HAVING**

***total\_flags*** > 500

**ORDER** **BY**

*hour* **DESC**, *event\_count* **DESC**

)

**SELECT** \* **FROM** *TrafficAnomalies\_dist*;