Group II

Relationship between

Crimes & Location

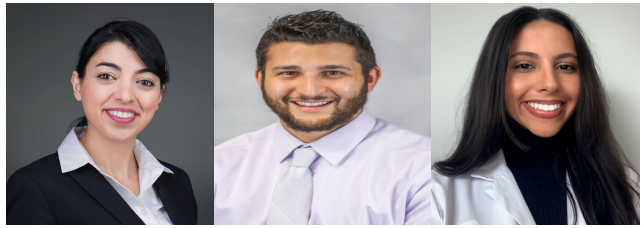
in

2021

in

Austin, Texas

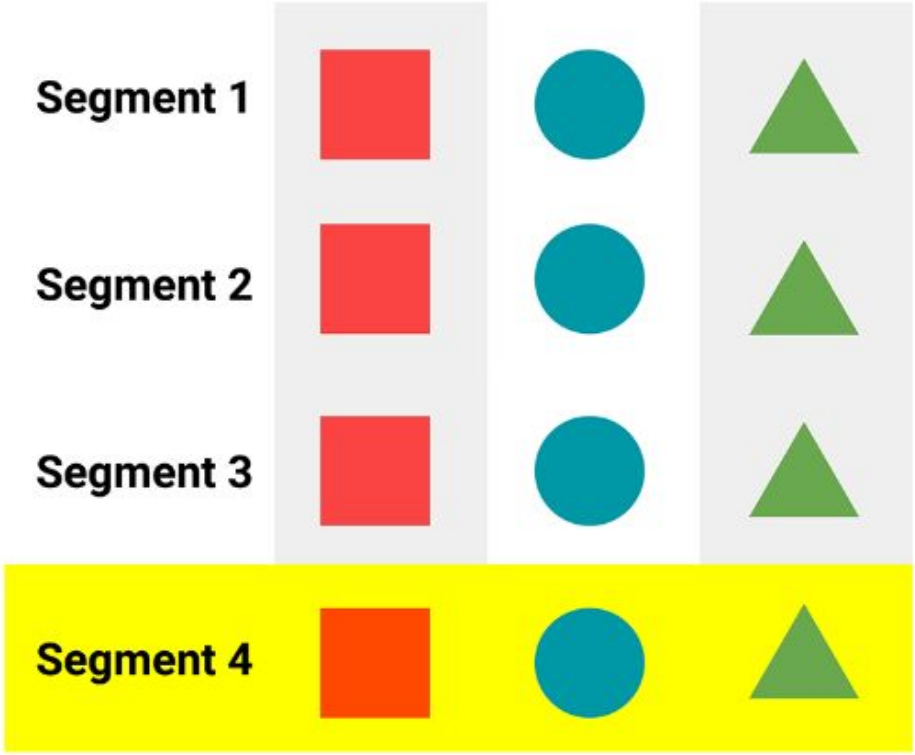
Team



Hasti: Consider the README

Dave: Visualization of the project

Diana: Machine Learning



Reason for this project

- Safer areas in Austin, Texas
- Dataset for 2021

Filter Data

	A	B	E	H	K	L	M	Y	Z
1	Incident Number	Highest Offense Description	Occurred Date Time	Report Date Time	Location T	Address	Zip Code	Latitude	Longitude
2	[illegible]5033598	PROTECTIVE ORDER	8/28/2018 7:23	8/28/2018 7:23	RESIDENC	UNKNOWN			
3	20195014472	BURGLARY OF VEHICLE	[illegible]	[illegible]	RESIDENC	3311 RED	78705	30.29225	-97.7258
4	2015150483	RAPE	12/1/2014 9:30	1/15/2015 9:30	FIELD / W(	00 BLOCK UNKNOWN			
5	20195044494	RAPE	10/26/2019 0:30	10/28/2019 15:27	OTHER / U	00 BLOCK UNKNOWN			
6	20191561862	POSSESSION OF MARIJUANA	6/5/2019 22:35	6/5/2019 22:35	RESIDENC	6201 SNEE	78744	30.19394	-97.7634
7	20191511203	EVADING / FOOT	5/31/2019 14:58	5/31/2019 14:58	RESIDENC	4618 MAN	78745	30.22573	-97.7913
8	2019990714	BURGLARY OF RESIDENCE	4/9/2019 12:06	4/9/2019 12:06	RESIDENC	2101 TERI	78744	30.20161	-97.752
9	20202411238	DISTURBANCE - OTHER	8/28/2020 18:50	8/28/2020 18:50	RESIDENC	3605 STECK AVE #1004			
10	20185031502	DAMAGE CITY PROP	8/15/2018 9:22	8/15/2018 9:22	RESIDENC	8011 E PARMER LN			
11	20205008273	PROTECTIVE ORDER	2/27/2020 10:10	2/27/2020 10:10	RESIDENC	UNKNOWN			
12	20191460782	FAMILY DISTURBANCE	5/26/2019 10:07	5/26/2019 10:07	RESIDENC	7500 S IH 3	78745	30.18167	-97.7774
13	20125010202	INDECENCY WITH A CHILD/CON	3/7/2012 12:27	3/7/2012 12:27	RESIDENC	2500 BLOC	78741		
14	20191420177	DATING DISTURBANCE	5/22/2019 3:09	5/22/2019 3:09	RESIDENC	3506 MAN	78704	30.23843	-97.7834
15	20081851973	ASSAULT W/INJURY-FAM/DATE	7/3/2008 20:52	7/3/2008 20:52	RESIDENC	4801 COPF	78744	30.19936	-97.7387
16	20205044863	PROTECTIVE ORDER	11/5/2020 11:19	11/5/2020 11:19	RESIDENC	UNKNOWN			
17	2011321905	RAPE	2/1/2011 19:51	2/1/2011 19:51	RESIDENC	1000 BLOC	78745		
18	2019791293	FAMILY DISTURBANCE	3/20/2019 17:23	3/20/2019 17:23	RESIDENC	1602 MAT	78745	30.19822	-97.8031
19	2019721051	BURGLARY OF RESIDENCE	3/13/2019 14:43	3/13/2019 14:43	RESIDENC	7302 SHAD	78745	30.18878	-97.788
20	20198003454	THEFT OF BICYCLE	9/10/2019 0:15	9/10/2019 14:41	RESIDENC	3506 SPEE	78705	30.30006	-97.7342
21	20195038708	FRAUD - OTHER	9/19/2019 12:00	9/19/2019 15:11	RESIDENC	3001 JACK	78723	30.30838	-97.6748
22	20192560087	FAMILY DISTURBANCE	9/13/2019 1:04	9/13/2019 1:27	RESIDENC	8038 EXCHANGE DR; BUILDING 9			
23	20192301136	AGG ASLT STRANGLE/SUFFOCAT	7/29/2019 0:00	8/18/2019 17:05	RESIDENC	1610 MORADO CIR			
24	20192570423	CRIMINAL TRESPASS	9/14/2019 9:58	9/14/2019 9:58	RESIDENC	5100 BLOC	78744		

Data loaded into python from pgadmin

In [51]:
```
url = "postgresql://postgres:password@database-1.cxzzhtb4x7ub.us-east-2.rds.amazonaws.com:5432/postgres"
```

In [52]:
```
import pandas as pd

from sqlalchemy import create_engine

engine = create_engine(url)

connect = engine.connect()
query = "select * from crime"

dataframe = pd.read_sql(query, con=connect)
dataframe
```

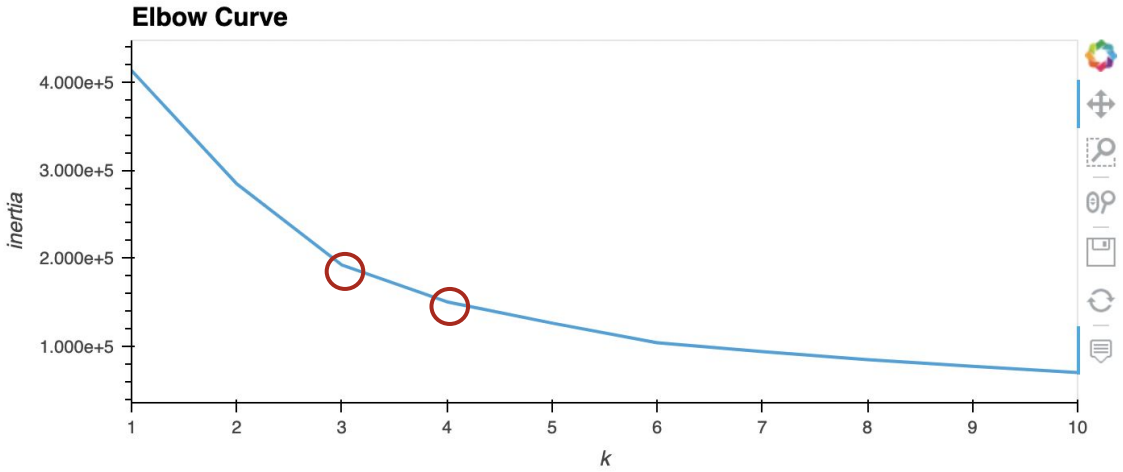
```
Exception during reset or similar
Traceback (most recent call last):
  File "C:\Users\lette\anaconda3\lib\site-packages\sqlalchemy\pool\base.py", line 682, in _finalize_fairy
    fairy._reset(pool)
  File "C:\Users\lette\anaconda3\lib\site-packages\sqlalchemy\pool\base.py", line 887, in _reset
    pool._dialect.do_rollback(self)
  File "C:\Users\lette\anaconda3\lib\site-packages\sqlalchemy\engine\default.py", line 667, in do_rollback
    dbapi_connection.rollback()
psycopg2.OperationalError: server closed the connection unexpectedly
        This probably means the server terminated abnormally
        before or while processing the request.
server closed the connection unexpectedly
        This probably means the server terminated abnormally
        before or while processing the request.
```

Out[52]:

	incident_number	highest_offense_description	occurred_date	occurred_time	report_date	report_time	zip_code	census_tract	clearance_date	x_coo
0	20213650197	ASSAULT W/INJURY-FAM/DATE VIOL	2021-12-31	436	2021-12-31	436	78736	330	2021-12-31	3
1	20213650632	ASSAULT W/INJURY-FAM/DATE VIOL	2021-12-31	1223	2021-12-31	1223	78726	346	2021-12-31	3
2	20213650144	AGG ASLT STRANGLE/SUFFOCATE	2021-12-31	244	2021-12-31	244	78752	433	2021-12-31	3

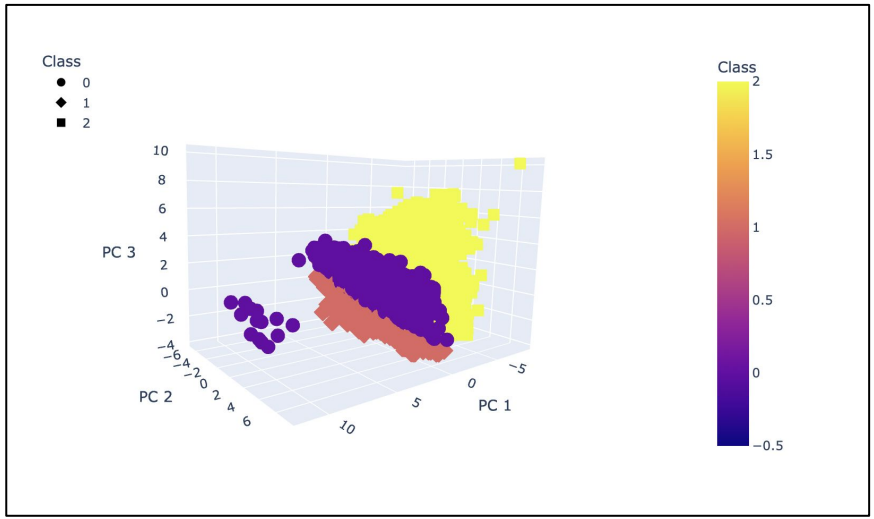
Introduction into Clustering Model

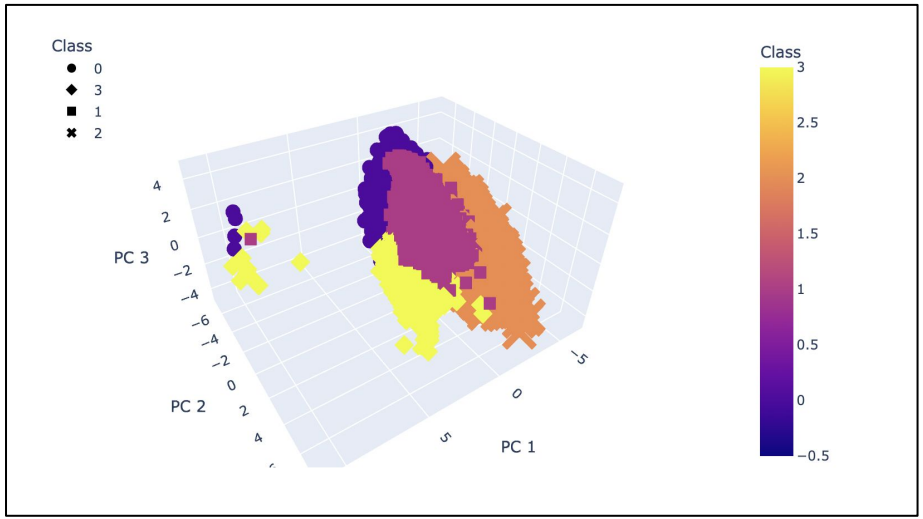
- Unsupervised clustering
  - Similar points cluster together while dissimilar points are separated
  - Census tract grouping vs. zip code grouping (geographical boundaries vs cluster of lines)
  - Creation of elbow curve in order to figure out how many clusters would be best (K-means)



K-means Model

- Creation of a 3D-Scatter K-means model
  - Visualization of clusters created
  - k = 3 vs k = 4?
  - PC1 = Highest Offense Description
  - PC2 = Longitude
  - PC3 = Latitude
- Reliable outputs?





Creation of a Neural Network

- Neural Network
  - Layers of neurons that individually perform computations
  - Recognition of pattern into quantitative outputs

- Keras Sequential Model
  - Where data flows from one layer to the next
  - Multiple dense layers that represent the input, hidden, and output layers
  - The total number of neurons and the activation function will be defined per each dense layer

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 80)                119840
_________________________________________________________________
dense_1 (Dense)              (None, 30)                2430
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 31
=================================================================
Total params: 122,301
Trainable params: 122,301
Non-trainable params: 0
```

Optimization Function vs Loss Metric

- Optimization Function
  - Shapes a neural network model
  - Ensures that the model is performing to its best ability

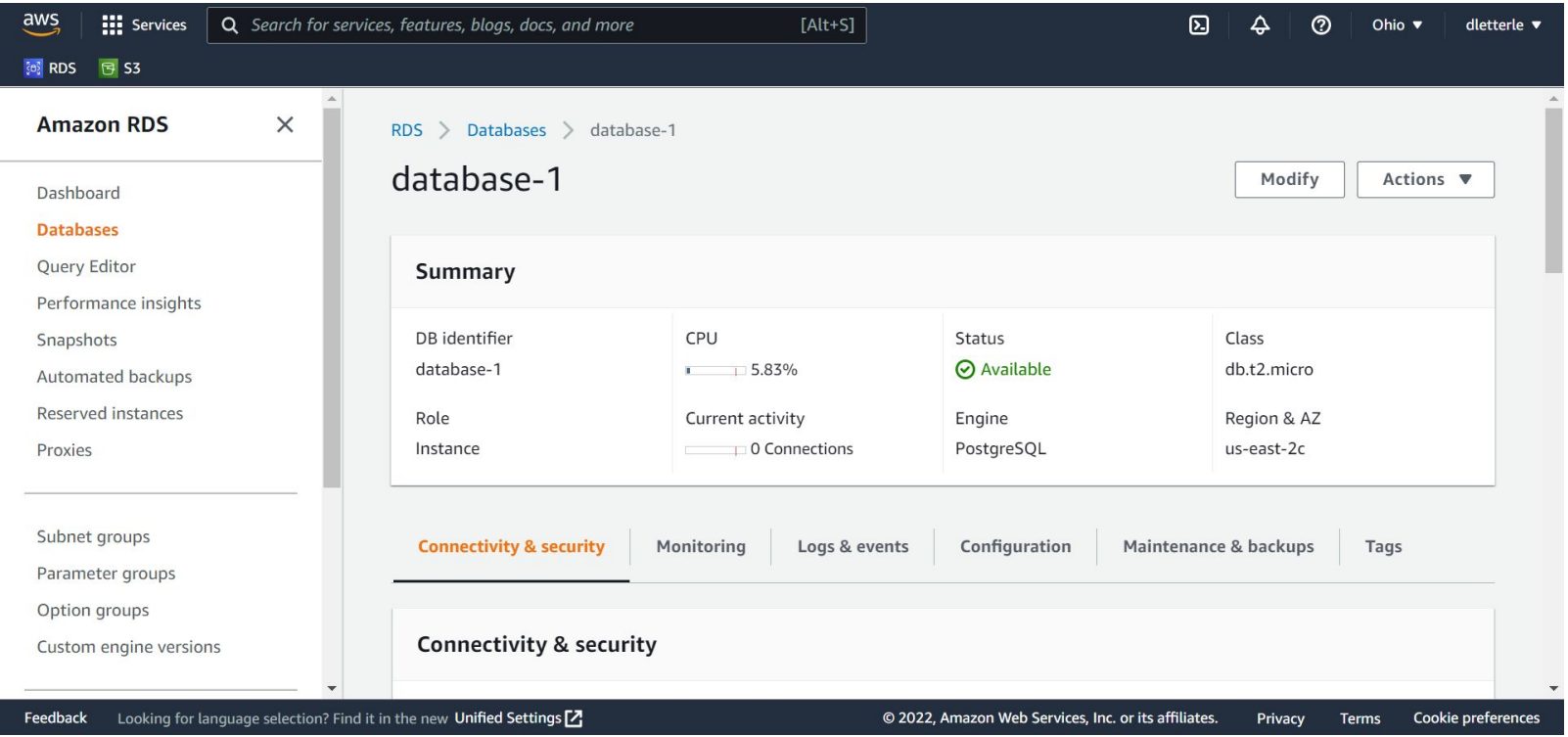
- Loss Metric
  - Tracks the model between each iteration and epoch
  - Evaluates the inaccuracy of each single output produced

```
Epoch 15/30
26275/26275 [==============================] - 15s 571us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 16/30
26275/26275 [==============================] - 15s 588us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 17/30
26275/26275 [==============================] - 17s 642us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 18/30
26275/26275 [==============================] - 15s 570us/sample - loss: 0.2973 - accuracy: 0.9122
Epoch 19/30
26275/26275 [==============================] - 15s 574us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 20/30
26275/26275 [==============================] - 15s 570us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 21/30
26275/26275 [==============================] - 16s 608us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 22/30
26275/26275 [==============================] - 15s 577us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 23/30
26275/26275 [==============================] - 16s 593us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 24/30
26275/26275 [==============================] - 16s 619us/sample - loss: 0.2973 - accuracy: 0.9122
Epoch 25/30
26275/26275 [==============================] - 18s 675us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 26/30
26275/26275 [==============================] - 16s 601us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 27/30
26275/26275 [==============================] - 15s 575us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 28/30
26275/26275 [==============================] - 15s 568us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 29/30
26275/26275 [==============================] - 15s 563us/sample - loss: 0.2974 - accuracy: 0.9122
Epoch 30/30
26275/26275 [==============================] - 14s 543us/sample - loss: 0.2974 - accuracy: 0.9122
```

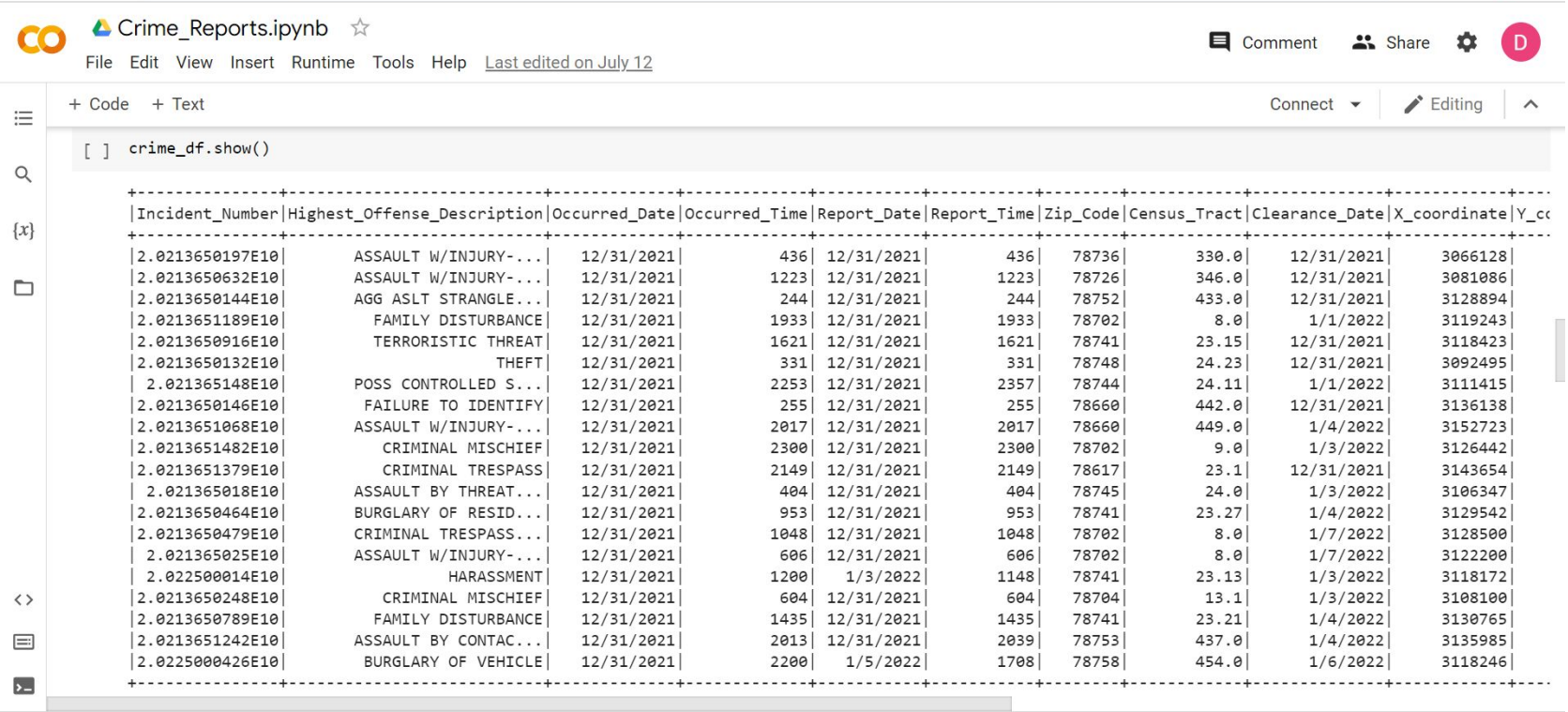
```
1  # Evaluate the model using the test data
2  model_loss, model_accuracy = nn.evaluate(X_test_scaled,y_test,verbose=2)
3  print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

```
8759/1 - 2s - loss: 0.2873 - accuracy: 0.9154
Loss: 0.25219930568736554, Accuracy: 0.9154012799263
```

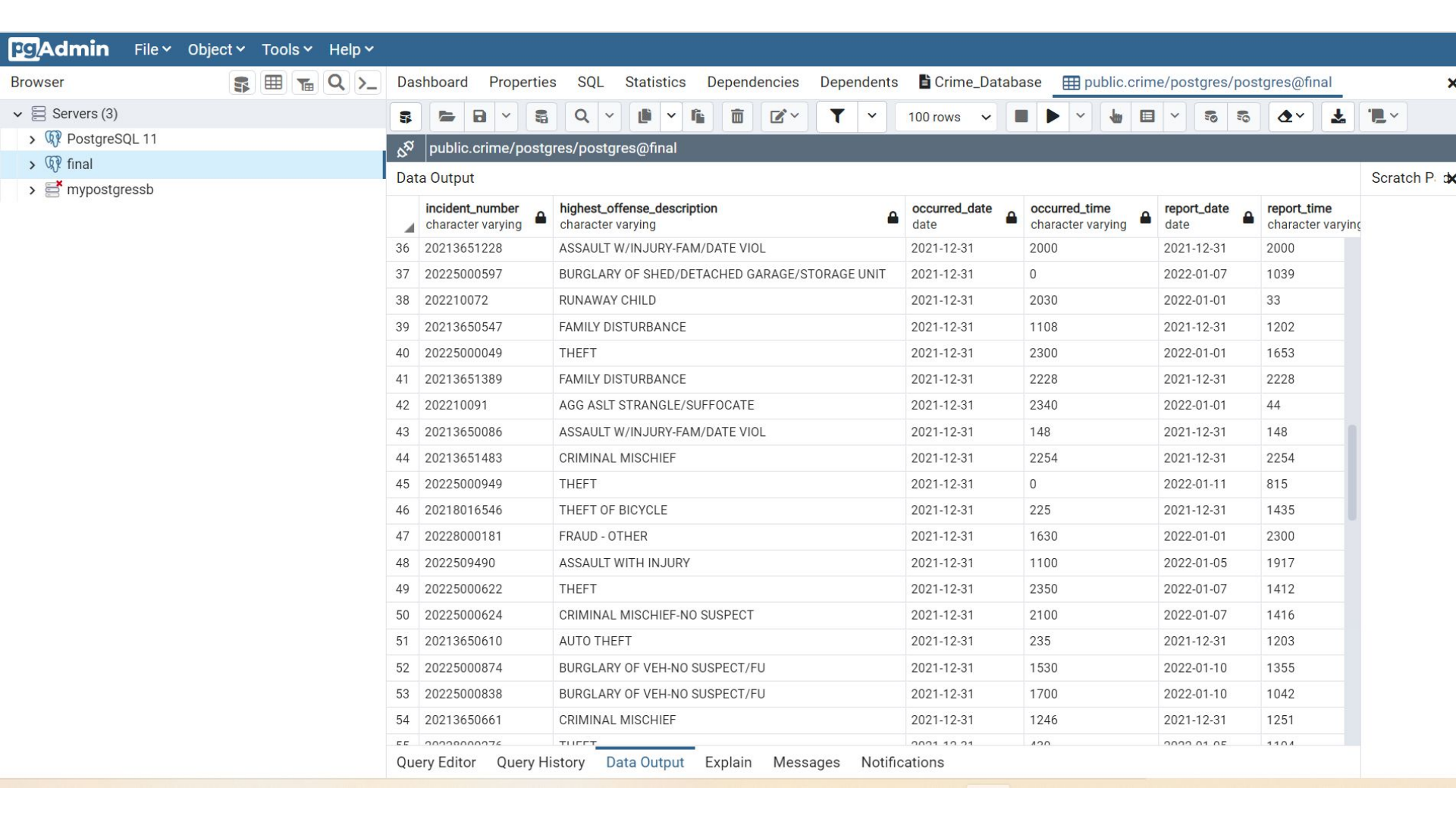
Data loaded into AWS and bucket was created



Data was reviewed in Google Colab and created database in pgadmin



Incident_Number	Highest_Offense_Description	Occurred_Date	Occurred_Time	Report_Date	Report_Time	Zip_Code	Census_Tract	Clearance_Date	X_coordinate	Y_cc
2.0213650197E10	ASSAULT W/INJURY-...	12/31/2021	436	12/31/2021	436	78736	330.0	12/31/2021	3066128	
2.0213650632E10	ASSAULT W/INJURY-...	12/31/2021	1223	12/31/2021	1223	78726	346.0	12/31/2021	3081086	
2.0213650144E10	AGG ASLT STRANGLE...	12/31/2021	244	12/31/2021	244	78752	433.0	12/31/2021	3128894	
2.0213651189E10	FAMILY DISTURBANCE	12/31/2021	1933	12/31/2021	1933	78702	8.0	1/1/2022	3119243	
2.0213650916E10	TERRORISTIC THREAT	12/31/2021	1621	12/31/2021	1621	78741	23.15	12/31/2021	3118423	
2.0213650132E10	THEFT	12/31/2021	331	12/31/2021	331	78748	24.23	12/31/2021	3092495	
2.021365148E10	POSS CONTROLLED S...	12/31/2021	2253	12/31/2021	2357	78744	24.11	1/1/2022	3111415	
2.0213650146E10	FAILURE TO IDENTIFY	12/31/2021	255	12/31/2021	255	78660	442.0	12/31/2021	3136138	
2.0213651068E10	ASSAULT W/INJURY-...	12/31/2021	2017	12/31/2021	2017	78660	449.0	1/4/2022	3152723	
2.0213651482E10	CRIMINAL MISCHIEF	12/31/2021	2300	12/31/2021	2300	78702	9.0	1/3/2022	3126442	
2.0213651379E10	CRIMINAL TRESPASS	12/31/2021	2149	12/31/2021	2149	78617	23.1	12/31/2021	3143654	
2.021365018E10	ASSAULT BY THREAT...	12/31/2021	404	12/31/2021	404	78745	24.0	1/3/2022	3106347	
2.0213650464E10	BURGLARY OF RESID...	12/31/2021	953	12/31/2021	953	78741	23.27	1/4/2022	3129542	
2.0213650479E10	CRIMINAL TRESPASS...	12/31/2021	1048	12/31/2021	1048	78702	8.0	1/7/2022	3128500	
2.021365025E10	ASSAULT W/INJURY-...	12/31/2021	606	12/31/2021	606	78702	8.0	1/7/2022	3122200	
2.022500014E10	HARASSMENT	12/31/2021	1200	1/3/2022	1148	78741	23.13	1/3/2022	3118172	
2.0213650248E10	CRIMINAL MISCHIEF	12/31/2021	604	12/31/2021	604	78704	13.1	1/3/2022	3108100	
2.0213650789E10	FAMILY DISTURBANCE	12/31/2021	1435	12/31/2021	1435	78741	23.21	1/4/2022	3130765	
2.0213651242E10	ASSAULT BY CONTAC...	12/31/2021	2013	12/31/2021	2039	78753	437.0	1/4/2022	3135985	
2.0225000426E10	BURGLARY OF VEHICLE	12/31/2021	2200	1/5/2022	1708	78758	454.0	1/6/2022	3118246	

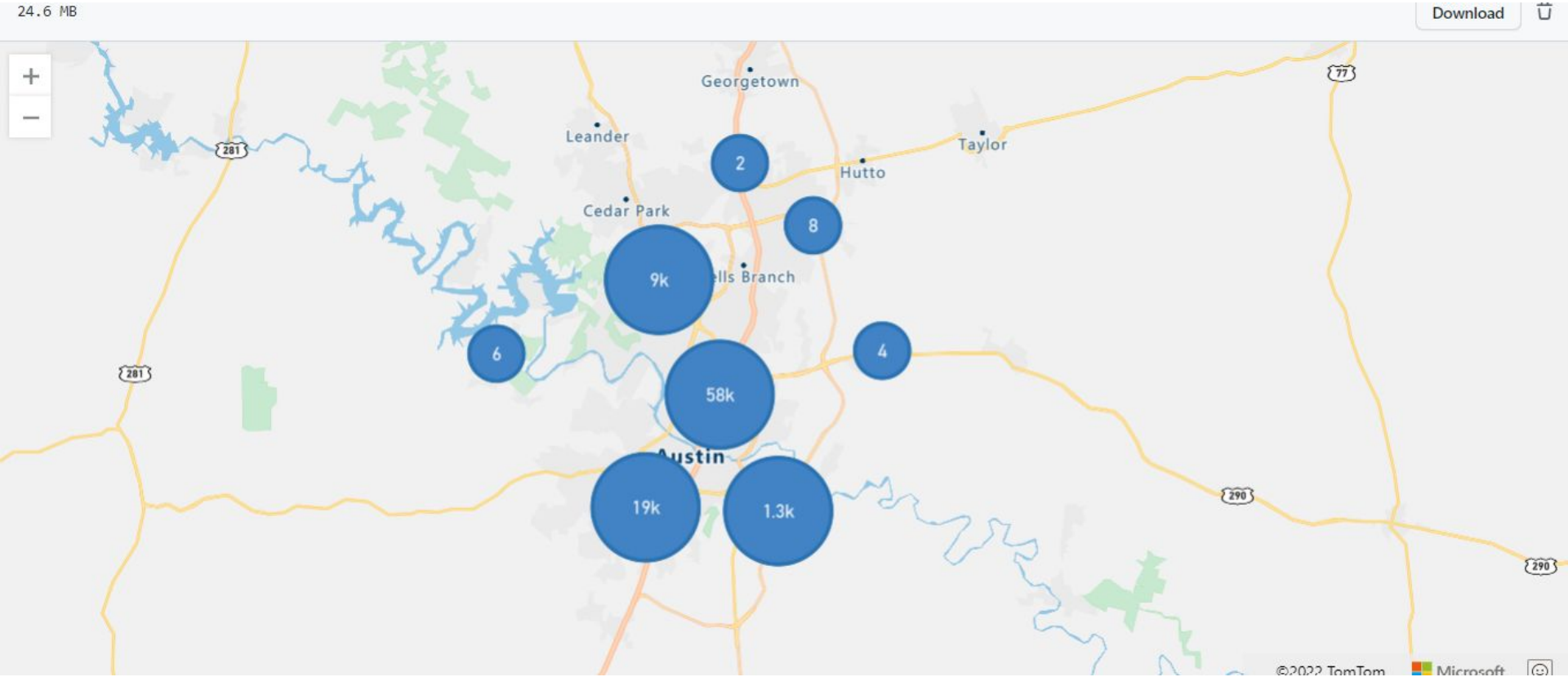

pgAdmin
File
Object
Tools
Help
Browser
Servers (3)
PostgreSQL 11
final
mypostgressb
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
Crime_Database
public.crime/postgres/postgres@final
100 rows
public.crime/postgres/postgres@final
Data Output
Scratch P
incident_number character varying
highest_offense_description character varying
occurred_date date
occurred_time character varying
report_date date
report_time character varying
36 20213651228 ASSAULT W/INJURY-FAM/DATE VIOL 2021-12-31 2000 2021-12-31 2000
37 20225000597 BURGLARY OF SHED/DETACHED GARAGE/STORAGE UNIT 2021-12-31 0 2022-01-07 1039
38 202210072 RUNAWAY CHILD 2021-12-31 2030 2022-01-01 33
39 20213650547 FAMILY DISTURBANCE 2021-12-31 1108 2021-12-31 1202
40 20225000049 THEFT 2021-12-31 2300 2022-01-01 1653
41 20213651389 FAMILY DISTURBANCE 2021-12-31 2228 2021-12-31 2228
42 202210091 AGG ASLT STRANGLE/SUFFOCATE 2021-12-31 2340 2022-01-01 44
43 20213650086 ASSAULT W/INJURY-FAM/DATE VIOL 2021-12-31 148 2021-12-31 148
44 20213651483 CRIMINAL MISCHIEF 2021-12-31 2254 2021-12-31 2254
45 20225000949 THEFT 2021-12-31 0 2022-01-11 815
46 20218016546 THEFT OF BICYCLE 2021-12-31 225 2021-12-31 1435
47 20228000181 FRAUD - OTHER 2021-12-31 1630 2022-01-01 2300
48 2022509490 ASSAULT WITH INJURY 2021-12-31 1100 2022-01-05 1917
49 20225000622 THEFT 2021-12-31 2350 2022-01-07 1412
50 20225000624 CRIMINAL MISCHIEF-NO SUSPECT 2021-12-31 2100 2022-01-07 1416
51 20213650610 AUTO THEFT 2021-12-31 235 2021-12-31 1203
52 20225000874 BURGLARY OF VEH-NO SUSPECT/FU 2021-12-31 1530 2022-01-10 1355
53 20225000838 BURGLARY OF VEH-NO SUSPECT/FU 2021-12-31 1700 2022-01-10 1042
54 20213650661 CRIMINAL MISCHIEF 2021-12-31 1246 2021-12-31 1251
Query Editor
Query History
Data Output
Explain
Messages
Notifications

Data transformed to geojson

logic.js | config.js | crime_updated.geojson

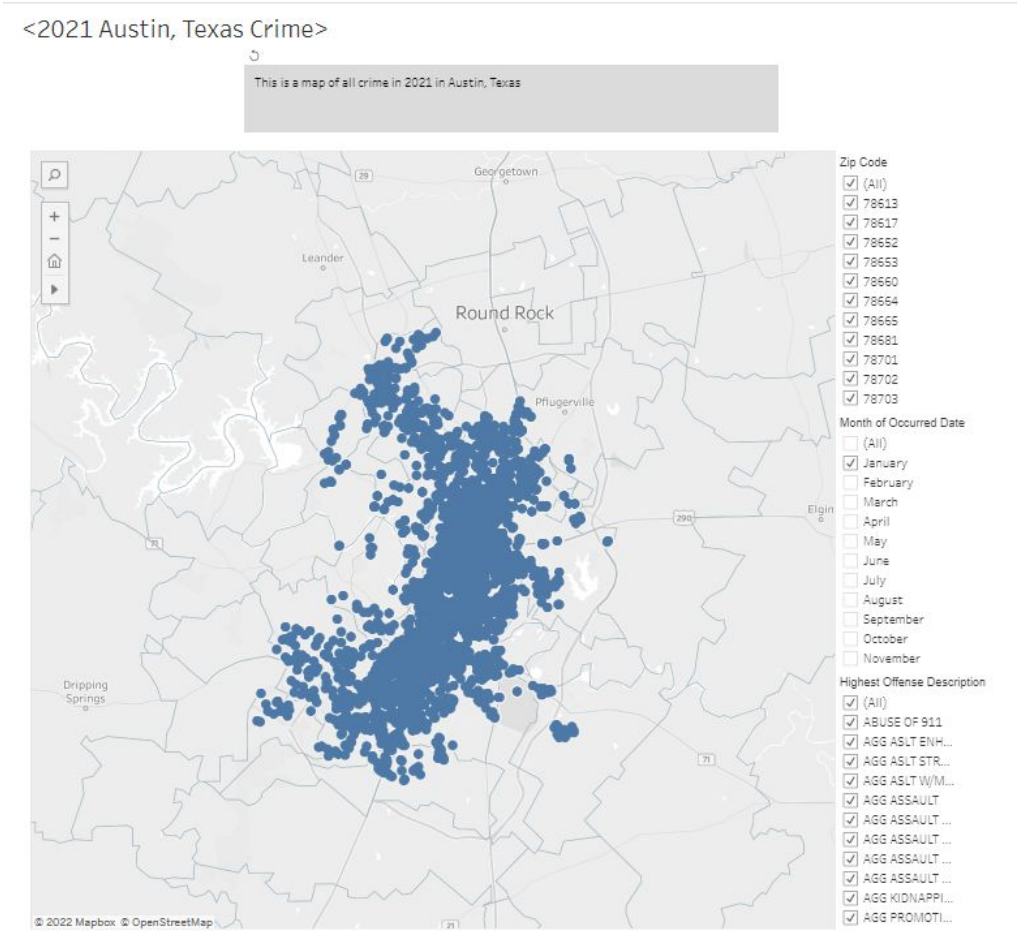

```
{
    "type": "FeatureCollection",
    "name": "crime_updated",
    "crs": {
        "type": "name",
        "properties": {
            "name": "urn:ogc:def:crs:OGC:1.3:CRS84"
        }
    },
    "features": [
        {
            "type": "Feature",
            "properties": {
                "Incident_Number": "20213650197",
                "Highest_Offense_Description": "ASSAULT W/INJURY-FAM/DATE VIOL",
                "Occurred_Date": "12/31/2021",
                "Occurred_Time": "436",
                "Report_Date": "12/31/2021",
                "Report_Time": "436",
                "Zip_Code": "78736",
                "Census_Tract": "330",
                "Clearance_Date": "12/31/2021",
                "X_Coordinate": 3066128.0,
                "Y_Coordinate": 3066128.0,
                "Latitude": 30.23640862,
                "Longitude": -97.89627861
            },
            "geometry": {
                "type": "Point",
                "coordinates": [
                    -97.89627861,
                    30.23640862
                ]
            }
        },
        {
```

GeoJson



https://github.com/Hastieiliat/Final-Project/blob/main/crime_updated.geojson

Tableau



https://public.tableau.com/app/profile/david8104/viz/Crime_16589473812840/Story1?publish=yes

Recommendations for future analysis

Analysis of a specific crime of interest over multiple years

Complete analysis of other years of data

Analysis of similar communities crime rates to that of Austin, Texas

FINISH

THANK YOU!