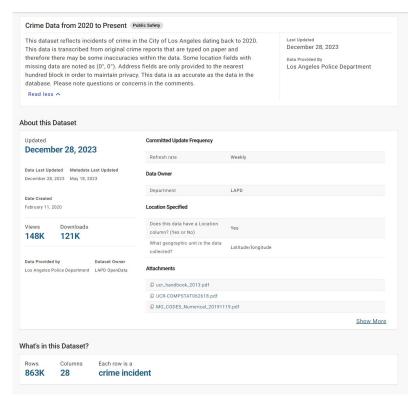


LOS ANGELES CRIME DATA ANALYTICS DASHBOARD

Los Angeles Crime Dataset: What is it?

- https://data.lacity.org/
- Los Angeles Police department
- Over 850k rows of crime incidents
- Updated weekly (2020 2023)



Los Angeles Crime Dataset: Target Audience

- LA Police Department
 - Hotspots
 - Trends
- Policy Makers
 - Preventative measures
 - Measure a policy's effectiveness
- LA Citizens
 - Aware of surroundings
 - Steps to ensure their safety







Los Angeles Crime Dataset: Extraction

- Using sodapy and pandas to extract raw data
- Results was then converted into a pandas dataframe for cleaning

```
import pandas as pd
from sodapy import Socrata

# Get my LA City App Token from the api_keys module file
from api_keys import lacity_app_token

# Reference: https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/about_data
client = Socrata("data.lacity.org", lacity_app_token)

# Approx: 7-8 Minutes
results = client.get_all("2nrs-mtv8")

# Convert to pandas DataFrame
crime_df = pd.DataFrame.from_records(results)
crime_df
```

Los Angeles Crime Dataset: Extraction

- Experimented with using JSON api
- Extraction approximately took 7-8 mins to download

```
create db ison.py > ...
   import requests
   import pandas as pd
   from pathlib import Path
   from datetime import datetime
   # Define the base URL and initial parameters
   base url = "https://data.lacity.org/resource/2nrs-mtv8.json"
   params = {
        "$limit": 50000, # Number of records per request
        "$offset": 0
                         # Initial offset
   print(base_url, params["$limit"])
   #create empty list
   all records = []
   while True:
       response = requests.get(base url, params=params)
       if response.status code == 200:
            data = response.json()
```

Los Angeles Crime Data: Raw Dataset (Pandas DataFrame)

	dr_no	date_rptd	date_occ	time_occ	area	area_name	rpt_dist_no	part_1_2	crm_cd	crm_cd_desc		status	status_desc	crm_cd_1	location	lat	lon	crm_cd_2	cross_street	crm_cd_3	crm_cd_4
0	010304468	2020-01- 08T00:00:00.000	2020-01- 08T00:00:00.000	2230		Southwest	0377		624	BATTERY - SIMPLE ASSAULT		AO	Adult Other	624	1100 W 39TH PL	34.0141	-118.2978	NaN	NaN	NaN	NaN
1	190101086	2020-01- 02T00:00:00.000	2020-01- 01T00:00:00.000	0330		Central	0163		624	BATTERY - SIMPLE ASSAULT		IC	Invest Cont	624	700 S HILL ST	34.0459	-118.2545	NaN	NaN	NaN	NaN
2	200110444	2020-04- 14T00:00:00.000	2020-02- 13T00:00:00.000	1200	01	Central	0155		845	SEX OFFENDER REGISTRANT OUT OF COMPLIANCE		AA	Adult Arrest	845	200 E 6TH ST	34.0448	-118.2474	NaN	NaN	NaN	NaN
3	191501505	2020-01- 01T00:00:00.000	2020-01- 01T00:00:00.000	1730	15	N Hollywood	1543		745	VANDALISM - MISDEAMEANOR (\$399 OR UNDER)		IC	Invest Cont	745	5400 CORTEEN PL	34.1685	-118.4019	998	NaN	NaN	NaN
4	191921269	2020-01- 01T00:00:00.000	2020-01- 01T00:00:00.000	0415	19	Mission	1998		740	VANDALISM - FELONY (\$400 & OVER, ALL CHURCH VA		IC	Invest Cont	740	14400 TITUS ST	34.2198	-118.4468	NaN	NaN	NaN	NaN
857653	231606525	2023-03- 22T00:00:00.000	2023-03- 22T00:00:00.000	1000	16	Foothill	1602		230	ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT		IC	Invest Cont	230	12800 FILMORE ST	34.279	-118.4116	NaN	NaN	NaN	NaN
857654	231210064	2023-04- 12T00:00:00.000	2023-04- 12T00:00:00.000	1630	12	77th Street	1239		230	ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT		IC	Invest Cont	230	6100 S VERMONT AV	33.9841	-118.2915	NaN	NaN	NaN	NaN
857655	230115220	2023-07- 02T00:00:00.000	2023-07- 01T00:00:00.000	0001	01	Central	0154		352	PICKPOCKET		IC	Invest Cont	352	500 S MAIN ST	34.0467	-118.2485	NaN	NaN	NaN	NaN
857656	230906458	2023-03- 05T00:00:00.000	2023-03- 05T00:00:00.000	0900	09	Van Nuys	0914		745	VANDALISM - MISDEAMEANOR (\$399 OR UNDER)		IC	Invest Cont	745	14500 HARTLAND ST	34.1951	-118.4487	NaN	NaN	NaN	NaN
857657	230319786	2023-11- 10T00:00:00.000	2023-11- 09T00:00:00.000	2300	03	Southwest	0395		331	THEFT FROM MOTOR VEHICLE - GRAND (\$950.01 AND		IC	Invest Cont	331	4100 S HOBART BL	34.0091	-118.3078	NaN	NaN	NaN	NaN
857658 rows × 28 columns																					

Los Angeles Crime Data: Cleaning The Dataset

- Jupyter Notebook
- Convert to the correct data types
- Removed unnecessary columns
- Replaced Codes to meaningful data (Eg: M = Male)
- Create new summarised Crime Category column

```
import numpy as np
   # Sex Code to Actual:
   # M - Male
         Female
   # X - Unknown
   replace sex code = {"M": "Male",
                       "F": "Female",
                       "X": "Unknown",
                       "-": None}
   final crime df = final crime df[~final crime df["vict_sex"].isin(["H"])]
   final crime df["vict sex"] = final crime df["vict sex"].replace({np.nan: None})
   final crime df["vict sex"] = final crime df["vict sex"].replace(replace sex code)
   unique_victim_sex = final_crime_df["vict_sex"].unique()
   print("Unique Victim Sex:")
   print()
   print(unique victim sex)
Unique Victim Sex:
['Female' 'Male' 'Unknown' None]
```

Los Angeles Crime Data: Adding 'Crime Categories'

UC

UCR REPORTING - Return A

(Based on date of reporting)

Part I - Violent Crimes

HOMICIDE 110 (Homicide)

113 (Manslaughter)

RAPE 121 (Rape)

122 (Attempted Rape)

815 (Sexual Penetration w/ Foreign Object)

820 (Oral Copulation)

821 (Sodomy)

ROBBERY 210 (Robbery)

220 (Robbery - attempted)

AGG. ASSAULTS 230 (ADW)

231 (ADW against LAPD Police Officer)

235 (Child beating)

DV* 236 (Spousal beating)

250 (Shots Fired)

251 (Shots fired inhabited dwelling)

761 (Brandishing) 926 (Train Wrecking)



```
rape codes = [121, 122, 815, 820, 821]
robbery_codes = [210, 220]
agg_assault_codes = [230, 231, 235]
domestic violence codes = [236, 250, 251, 761, 926,
                          626, 627, 647, 763, 928, 930]
simple_assault_codes = [435, 436, 437, 622, 623, 624, 625]
burglary codes = [310, 320]
gta_codes = [510, 520, 433]
btfv_codes = [330, 331, 410, 420, 421]
personal theft codes = [350, 351, 352, 353, 450, 451, 452, 453]
other theft codes = [341, 343, 345, 440, 441, 442, 443, 444, 445,
                     470, 471, 472, 473, 474, 475, 480, 485, 487, 491]
ucr dict = {"Homicide": homicide codes.
            "Rape": rape codes,
            "Robbery": robbery codes,
            "Aggravated Assault": agg assault codes,
            "Domestic Violence": domestic_violence_codes,
            "Simple Assault": simple assault codes,
            "Burglary": burglary_codes,
            "Grand Theft Auto": gta_codes,
            "Personal Theft": personal theft codes,
            "Other Theft": other_theft_codes}
```

Los Angeles Crime Data: Final Dataset

Reduced to:

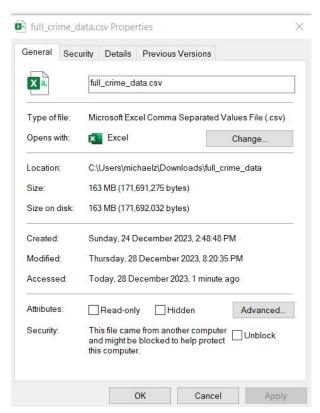
- 16 Columns
- 800k records
- Converted to appropriate data types
- New Crime Category field
- Translated:
- Descent Code
- Sex Code

```
dr_no
date_rptd
date occ
time_occ
area_name
crime category
crm cd
crm cd desc
vict age
vict sex
vict descent
premis_desc
location
cross_street
lat
lon
```

Los Angeles Crime Data: Final Dataset

Original raw csv file size: 250MB

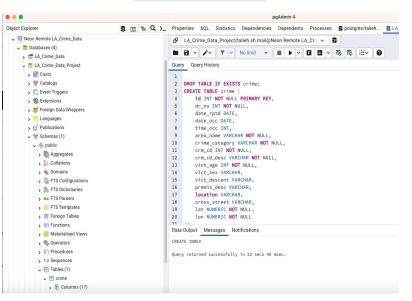
Once cleaned: 163MB



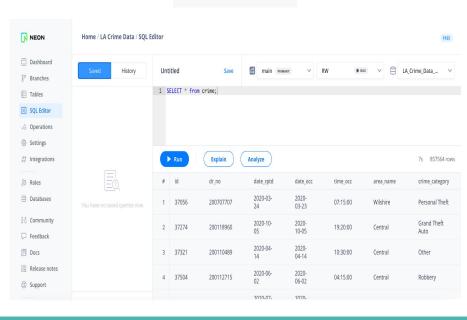
PostgreSQL Database: Creation & Cloud Hosting











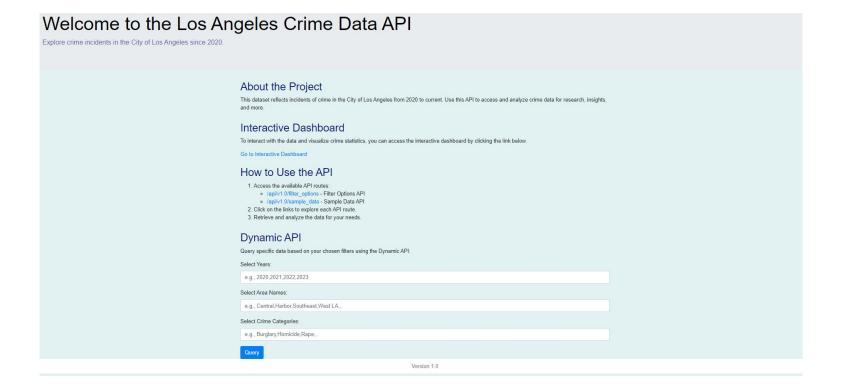
Flask Web Application: Crime Dataset API



- SQL Database Connection
- Data Querying and Response
- Route Definitions:
 - Homepage
 - Interactive Dashboard
 - API Endpoints

```
app render.pv M X
                                                                  D ∨ 11 €0
app_render.py >  filter_options
     # Flask API App Setup
     app = Flask(__name__)
     # Configure JSONIFY SETTINGS to control JSON serialization behavior
     app.config['JSONIFY_PRETTYPRINT_REGULAR'] = False # Disable pretty-printing JSON responses
     *************************************
     @app.route("/")
        # Run the homepage HTML webpage (homepage.html)
        return render template("homepage.html")
     ######## Route #2 (Interactive Dashboard) #############
     @app.route("/frontend")
     def frontend():
        # Run the Analytical Dashboard HTML webpage (index.html)
        return render_template("index.html")
     @app.route("/api/v1.0/filter options")
     def filter options():
        session = Session(engine)
         query_years = session.query(func.extract('year', crime.date_occ).\
                               label('vear')).distinct().all()
```

Flask Web Application: Crime Dataset API Landing Page

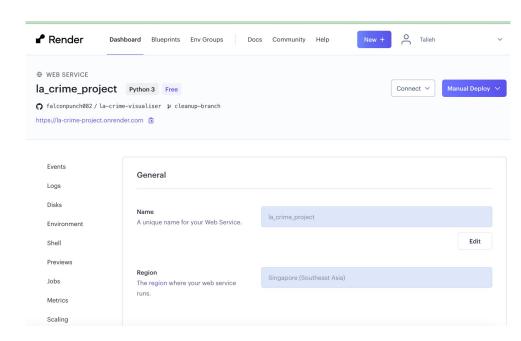




Flask Application: Cloud Hosting with Render

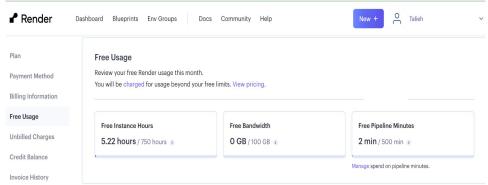
Why Render?

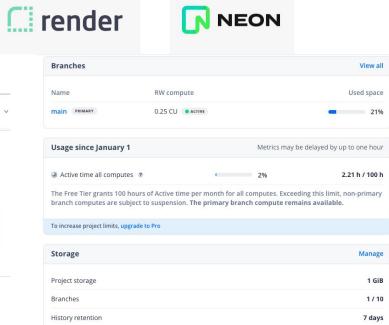
- Automated Deployment
- Ease of Use
- Automatic Scaling
- Developer-Friendly
- Cost-Effective



Cloud Hosting: Challenges & Solutions

- Database hosting on cloud
- Web application hosting on cloud





658 MiB / 3 GiB

Current data size for main PRIMARY

Interactive Filters: Populating For All Options











Dynamic API Route: Query Filtered Data from PostgreSQL Table



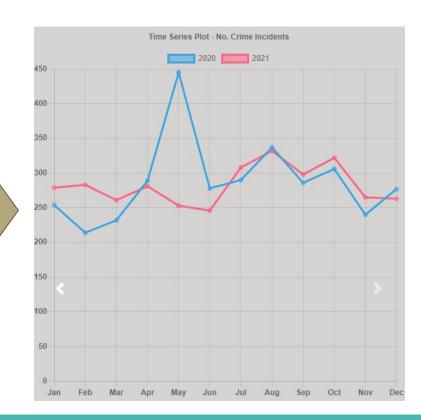
@app.route("/api/v1.0/<years_str>/<area_names_str>/<crime_categories_str>")

```
"vears": [
 2020,
 2021
"area names": [
  "Central"
"crime categories": [
  "Burglary"
"crime data": [
   "id": 619,
   "dr no": 200105263,
   "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
"date_occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00".
    "area_name": "Central",
    "crime_category": "Burglary",
    "crm cd": 310.
    "crm cd desc": "BURGLARY".
    "vict_age": 36,
"vict_sex": "Female",
    "vict descent": "White".
     "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
     "location": "700 S BROADWAY",
    "cross_street": null,
    "lat": "34.0452",
    "lon": "-118.2534"
```

Visualisation #1: Time Series Plot (Chart.js)

```
"years": [
 2020.
 2021
'area names": [
 "Central"
crime categories": [
  "Burglary"
],
"crime_data": [
    "id": 619,
    "dr no": 200105263.
    "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
    "date occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00",
    "area name": "Central",
    "crime_category": "Burglary",
    "crm_cd": 310,
"crm_cd desc": "BURGLARY",
    "vict_age": 36,
    "vict_sex": "Female",
    "vict descent": "White",
    "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
    "location": "700 S BROADWAY".
    "cross_street": null,
    "lat": "34.0452",
"lon": "-118.2534"
    "id": 687,
    "dr no": 200105370,
    "date rptd": "Mon, 20 Jan 2020 00:00:00 GMT",
    "date_occ": "Mon, 20 Jan 2020 00:00:00 GMT",
    "time_occ": "17:00:00", 
"area_name": "Central",
    "crime category": "Burglary",
    "crm_cd": 310,
    "crm cd desc": "BURGLARY",
    "vict_age": 30,
    "vict sex": "Female".
    "vict_descent": "Other",
    "premis desc": "SINGLE FAMILY DWELLING",
    "location": "800 ALPINE
                                                        ST",
    "cross street": null.
    "lat": "34.0637",
    "lon": "-118.2440"
```

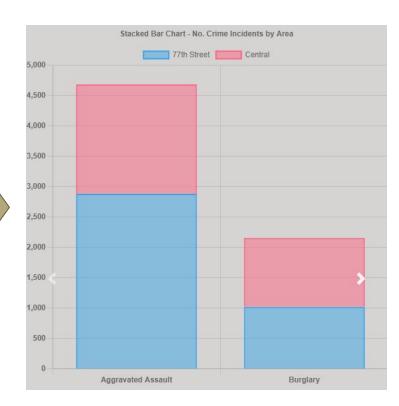
```
const init TimeSeries = (thisDataset) => {
   const monthLabels = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun
   // From the Oueried JSON Dataset, store the uniquely select
   const uniqueYearsSet = new Set(thisDataset.years);
   const uniqueYearsArray = [...uniqueYearsSet];
   const minYear = Math.min(...uniqueYearsArray);
   const maxYear = Math.max(...uniqueYearsArray);
   // Length of array determined by difference between max as
   const finalYears = Array.from( | length: maxYear - minYear
   const tsData = finalYears.map(Year => {
       const currentYearData = thisDataset.crime data.filter(i
       const getMonthCounts = Array.from({ length: 12 }, ( , m
         const currentMonth = monthIndex + 1;
         return currentYearData.filter(item => new Date(item.d
```



Visualisation #2: Stacked Bar Chart (Chart.js)

```
"years": [
 2020.
 2021
 'area names": [
  "Central"
"crime categories": [
  "Burglary"
],
"crime_data": [
    "id": 619,
    "dr no": 200105263.
    "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
    "date occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00",
    "area name": "Central",
    "crime_category": "Burglary",
    "crm_cd": 310,
"crm_cd desc": "BURGLARY",
    "vict age": 36,
    "vict_sex": "Female",
    "vict descent": "White",
    "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
    "location": "700 S BROADWAY",
    "cross_street": null,
    "lat": "34.0452",
"lon": "-118.2534"
    "id": 687,
    "dr no": 200105370,
    "date rptd": "Mon, 20 Jan 2020 00:00:00 GMT",
    "date_occ": "Mon, 20 Jan 2020 00:00:00 GMT",
    "time_occ": "17:00:00", 
"area_name": "Central",
    "crime category": "Burglary",
    "crm_cd": 310,
    "crm cd desc": "BURGLARY",
    "vict_age": 30,
    "vict sex": "Female",
    "vict_descent": "Other",
    "premis desc": "SINGLE FAMILY DWELLING",
    "location": "800 ALPINE
"cross street": null,
                                                          ST",
    "lat": "34.0637".
    "lon": "-118.2440"
```

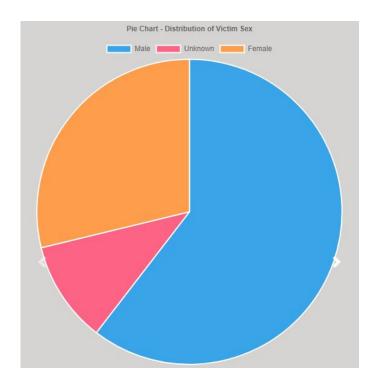
```
const init StackedBarChart = (thisDataset) => {
   const uniqueCrimes = thisDataset.crime categories:
   const uniqueAreas = thisDataset.area names;
   let barData = {
       labels: uniqueCrimes.
       datasets:
   uniqueAreas.forEach(Area => {
      let columnData = {
           label: Area,
           data: [],
                           // Dataset = count of Area Nam
          borderWidth: 2 // Column Width = 2
       uniqueCrimes.forEach(Crime => {
           let getCount = thisDataset.crime data.filter(it
          columnData.data.push(getCount);
      barData.datasets.push(columnData):
```



Visualisation #3: Pie Chart (Chart.js)

```
"years": [
 2020,
 2021
'area names": [
 "Central"
crime categories": [
  "Burglary"
],
"crime_data": [
    "dr no": 200105263.
    "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
    "date occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00",
    "area name": "Central",
    "crime_category": "Burglary",
    "crm_cd": 310,
"crm_cd desc": "BURGLARY",
    "vict age": 36,
    "vict_sex": "Female",
    "vict descent": "White",
    "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
    "location": "700 S BROADWAY",
    "cross_street": null,
    "lat": "34.0452",
"lon": "-118.2534"
    "id": 687,
    "dr no": 200105370,
    "date rptd": "Mon, 20 Jan 2020 00:00:00 GMT",
    "date_occ": "Mon, 20 Jan 2020 00:00:00 GMT",
    "time_occ": "17:00:00", 
"area_name": "Central",
    "crime category": "Burglary",
    "crm_cd": 310,
    "crm cd desc": "BURGLARY",
    "vict_age": 30,
    "vict sex": "Female".
    "vict_descent": "Other",
    "premis desc": "SINGLE FAMILY DWELLING",
    "location": "800 ALPINE
"cross street": null,
                                                         ST",
    "lat": "34.0637",
    "lon": "-118.2440"
```

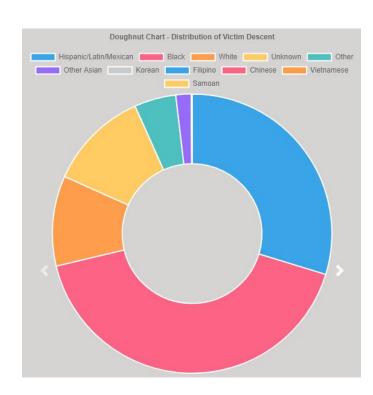
```
const init PieChart = (thisDataset) => {
   const nullValue = null:
   let uniqueSexes = [...new Set(thisDataset.crime data.map(ite
   uniqueSexes = uniqueSexes.filter(item => item !== nullValue
   let sliceData = []
   uniqueSexes.forEach(Sex => 4
       const getCount = thisDataset.crime data.filter(item =>
       sliceData.push(getCount);
   d3.select("#chartContainer3").html("");
   d3.select("#chartContainer3").html("<canvas id='jsChart Pie
   const pieCanvas = document.getElementById("jsChart_Pie").ge
   const pieChart = new Chart(pieCanvas. {
       type: 'pie'.
           labels: uniqueSexes,
           datasets: [{
               data: sliceData}]
```



Visualisation #4: Doughnut Chart (Chart.js)

```
"years": [
 2020.
 2021
 'area names": [
  "Central"
 crime categories": [
  "Burglary"
],
<u>"crime_</u>data": [
    "dr no": 200105263.
    "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
    "date occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00",
    "area name": "Central",
    "crime_category": "Burglary",
    "crm_cd": 310,
"crm_cd desc": "BURGLARY",
    "vict age": 36,
    "vict_sex": "Female",
    "vict descent": "White",
    "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
    "location": "700 S BROADWAY",
    "cross_street": null,
    "lat": "34.0452",
"lon": "-118.2534"
    "id": 687,
    "dr no": 200105370,
    "date rptd": "Mon, 20 Jan 2020 00:00:00 GMT",
    "date_occ": "Mon, 20 Jan 2020 00:00:00 GMT",
    "time_occ": "17:00:00", 
"area_name": "Central",
    "crime category": "Burglary",
    "crm_cd": 310,
    "crm cd desc": "BURGLARY",
    "vict_age": 30,
    "vict sex": "Female".
    "vict_descent": "Other",
     "premis desc": "SINGLE FAMILY DWELLING",
     "location": "800 ALPINE
                                                         ST",
    "cross street": null.
    "lat": "34.0637".
    "lon": "-118.2440"
```

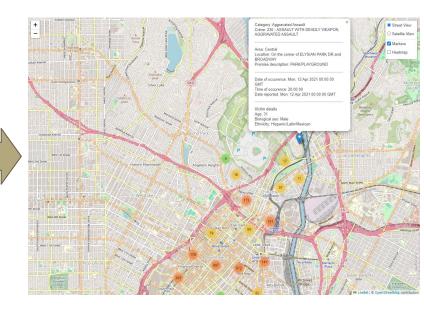
```
const init DonutChart = (thisDataset) => {
   const nullValue = null;
   let uniqueDescents = [...new Set(thisDataset.crime_data.ma
   uniqueDescents = uniqueDescents.filter(item => item !== no
    let sliceData = []
   uniqueDescents.forEach(Descent => {
       const getCount = thisDataset.crime data.filter(item =
       sliceData.push(getCount):
   d3.select("#chartContainer4").html("");
   d3.select("#chartContainer4").html("<canvas id='isChart Do
   const donutCanvas = document.getElementBvId("isChart Donu
   const donutChart = new Chart(donutCanvas, {
       type: 'doughnut'.
       data: {
```



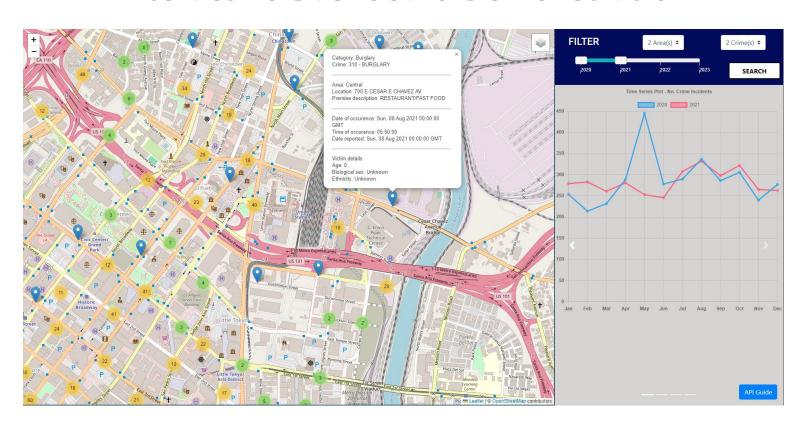
Visualisation #5: Interactive Map (Leaflet.js)

```
"years": [
 2020.
 2021
 'area names": [
  "Central"
 crime categories": [
  "Burglary
],
"<u>cri</u>me_data": [
    "dr no": 200105263.
    "date_rptd": "Sun, 19 Jan 2020 00:00:00 GMT",
    "date occ": "Sat, 18 Jan 2020 00:00:00 GMT",
    "time occ": "19:00:00",
    "area name": "Central",
    "crime_category": "Burglary",
    "crm_cd": 310,
"crm_cd desc": "BURGLARY",
    "vict age": 36,
    "vict_sex": "Female",
    "vict descent": "White",
    "premis desc": "MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)",
    "location": "700 S BROADWAY",
    "cross_street": null,
    "lat": "34.0452",
"lon": "-118.2534"
    "id": 687,
    "dr no": 200105370,
    "date rptd": "Mon, 20 Jan 2020 00:00:00 GMT",
    "date_occ": "Mon, 20 Jan 2020 00:00:00 GMT",
    "time_occ": "17:00:00", 
"area_name": "Central",
    "crime category": "Burglary",
    "crm_cd": 310,
    "crm cd desc": "BURGLARY",
    "vict_age": 30,
    "vict sex": "Female".
    "vict_descent": "Other",
    "premis_desc": "SINGLE FAMILY DWELLING",
    "location": "800 ALPINE
"cross street": null,
                                                          ST",
    "lat": "34.0637",
    "lon": "-118.2440"
```

```
const init Map = (thisDataset) => {
       resultData = thisDataset.crime data;
       let marks = L.markerClusterGroup();
       let hm pts = []:
       for(i = 0; i < resultData.length; i++) {</pre>
           let datapoint = resultData[i];
           let coords = [datapoint["lat"], datapoint["lon"]];
           let area = datapoint["area name"]:
           let category = datapoint["crime category"];
           let code = datapoint["crm cd"]:
           let descp = datapoint["crm cd desc"]:
           let street = datapoint["location"]:
           let x str = datapoint["cross street"];
           let occ date = datapoint["date occ"];
           let occ time = datapoint["time occ"]
           let rep = datapoint["date rptd"]:
           let premise = datapoint["premis desc"];
           let age = datapoint["vict age"]:
           let sex = datapoint["vict sex"];
           let descent = datapoint["vict descent"]:
```



Interactive Dashboard Demonstration



Thanks For Listening!

References

- <u>https://github.com/falconpunch082/la-crime-visualiser?tab=readme-ov-file#target-audience</u>
- https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/about data