

-----Portfolio Project 03-----

-----USA Airports Spatial Data Analysis-----

Import Basic Libraries

```
In [68]: ➤ import geopandas as gbd
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [14]: ➤ from shapely.geometry import Point
```

Load the Data Set

```
In [24]: ➤ airport_data = pd.read_csv('D:/QGIS/QGIS dataset/QGIS 3d/RainFall data set (Interpolation
```

```
In [6]: ➤ airport_data.head()
```

Out[6]:

	IATA	AIRPORT	CITY	STATE	COUNTRY	LATITUDE	LONGITUDE
0	ABQ	Albuquerque International	Albuquerque	NM	USA	35.040222	-106.609194
1	ANC	Ted Stevens Anchorage International	Anchorage	AK	USA	61.174320	-149.996186
2	ATL	William B Hartsfield-Atlanta Intl	Atlanta	GA	USA	33.640444	-84.426944
3	AUS	Austin-Bergstrom International	Austin	TX	USA	30.194533	-97.669872
4	BDL	Bradley International	Windsor Locks	CT	USA	41.938874	-72.683228

Data Preprocessing/ Feature Engineering

```
In [8]: ➤ airport_data.shape
```

Out[8]: (340, 7)

```
In [ ]: ➤ # For Numeric Data
```

```
In [81]: ➤ airport_data.describe().T
```

Out[81]:

	count	mean	std	min	25%	50%	75%	max
LATITUDE	340.0	38.853527	8.447379	13.483450	33.65784	38.998547	43.113107	71.285448
LONGITUDE	340.0	-98.462649	21.234125	-176.646031	-111.12648	-93.743140	-82.840772	-64.798556

```
In [ ]: # For Non-numeric Data
```

```
In [82]: airport_data.describe(include='object').T
```

Out[82]:

	count	unique	top	freq
IATA	340	340	ABQ	1
AIRPORT	340	340	Albuquerque International	1
CITY	337	324	Portland	2
state_code	337	54	CA	26
COUNTRY	340	1	USA	340

```
In [9]: airport_data.columns
```

Out[9]: Index(['IATA', 'AIRPORT', 'CITY', 'STATE', 'COUNTRY', 'LATITUDE', 'LONGITUDE'], dtype='object')

```
In [21]: airport_data.isnull().sum()
```

Out[21]: IATA 0
AIRPORT 0
CITY 3
STATE 3
COUNTRY 0
LATITUDE 0
LONGITUDE 0
dtype: int64

```
In [17]: type(geometry)
```

Out[17]: list

Importing the state ESRI ShapeFile of the USA

```
In [28]: us_states = gbd.read_file('D:/QGIS/QGIS dataset/QGIS 3d/RainFall data set (Interpolation
```

Import 2ns CSV file

```
In [29]: airport_data = pd.read_csv('D:/QGIS/QGIS dataset/QGIS 3d/RainFall data set (Interpolation
```

```
In [30]: airport_data.head()
```

Out[30]:

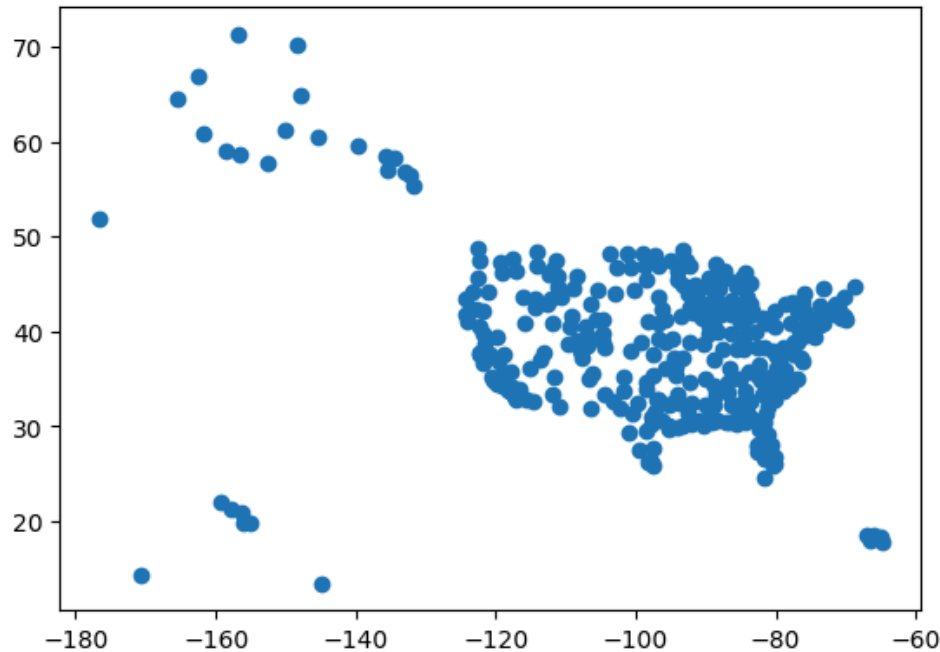
	IATA	AIRPORT	CITY	STATE	COUNTRY	LATITUDE	LONGITUDE
0	ABQ	Albuquerque International	Albuquerque	NM	USA	35.040222	-106.609194
1	ANC	Ted Stevens Anchorage International	Anchorage	AK	USA	61.174320	-149.996186
2	ATL	William B Hartsfield-Atlanta Intl	Atlanta	GA	USA	33.640444	-84.426944
3	AUS	Austin-Bergstrom International	Austin	TX	USA	30.194533	-97.669872
4	BDL	Bradley International	Windsor Locks	CT	USA	41.938874	-72.683228

Data Preprocessing For Spatial Data Analysis

```
In [47]: geometry = [Point(xy) for xy in zip(airport_data['LONGITUDE'],airport_data['LATITUDE'])]
```

```
In [49]: airport_us= gbd.GeoDataFrame(airport_data, geometry = geometry, crs=us_states.crs)
```

```
In [50]: airport_data.plot()  
plt.show()
```



```
In [51]: airport_data.head()
```

Out[51]:

	IATA	AIRPORT	CITY	state_code	COUNTRY	LATITUDE	LONGITUDE	geometry
0	ABQ	Albuquerque International	Albuquerque	NM	USA	35.040222	-106.609194	POINT (-106.60919 35.04022)
1	ANC	Ted Stevens Anchorage International	Anchorage	AK	USA	61.174320	-149.996186	POINT (-149.99619 61.17432)
2	ATL	William B Hartsfield-Atlanta Intl	Atlanta	GA	USA	33.640444	-84.426944	POINT (-84.42694 33.64044)
3	AUS	Austin-Bergstrom International	Austin	TX	USA	30.194533	-97.669872	POINT (-97.66987 30.19453)
4	BDL	Bradley International	Windsor Locks	CT	USA	41.938874	-72.683228	POINT (-72.68323 41.93887)

Attribute Joins

```
In [36]: # Importing the CSV which consists of state names and codes
```

```
In [52]: state_names_codes = pd.read_csv('D:/QGIS/QGIS dataset/QGIS 3d/RainFall data set (Interpol
```

```
In [53]: state_names_codes.head(10)
```

Out[53]:

	state_name	state_code
0	Alabama	AL
1	Alaska	AK
2	Arizona	AZ
3	Arkansas	AR
4	California	CA
5	Colorado	CO
6	Connecticut	CT
7	Delaware	DE
8	District of Columbia	DC
9	Florida	FL

```
In [40]: # Renaming the column heading
```

```
In [41]: airport_data.rename(columns={'STATE':'state_code'}, inplace=True)
```

```
In [42]: airport_data.columns
```

Out[42]: Index(['IATA', 'AIRPORT', 'CITY', 'state_code', 'COUNTRY', 'LATITUDE',
'LONGITUDE', 'geometry'],
dtype='object')

```
In [43]: # Join attributes
```

```
In [54]: airport_us = airport_us.merge(state_names_codes, on= 'state_code')
```

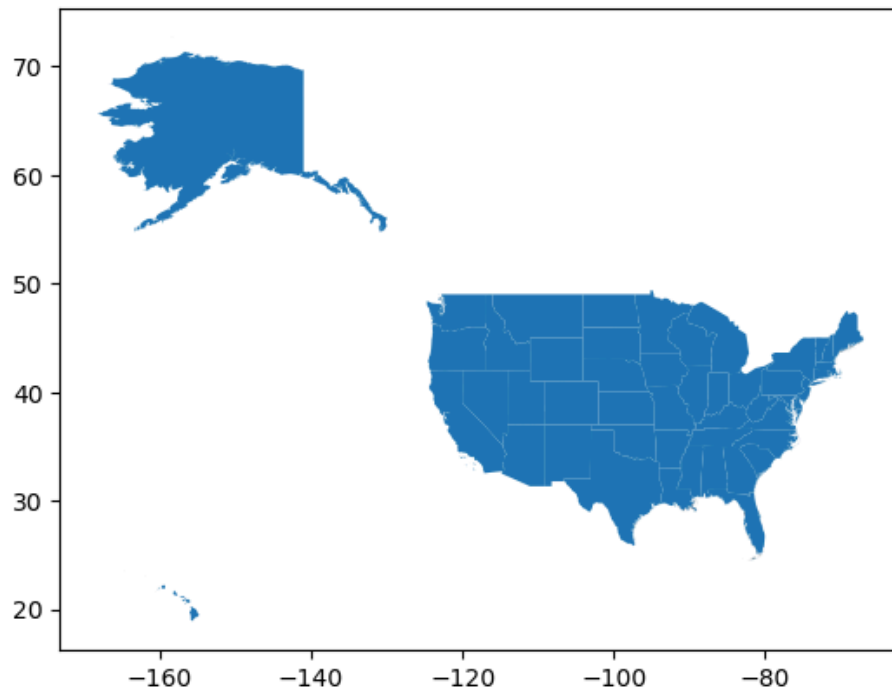
```
In [55]: airport_us.head()
```

Out[55]:

	IATA	AIRPORT	CITY	state_code	COUNTRY	LATITUDE	LONGITUDE	geometry	state_name
0	ABQ	Albuquerque International	Albuquerque	NM	USA	35.040222	-106.609194	POINT (-106.60919 35.04022)	New Mexico
1	HOB	Lea County Regional	Hobbs	NM	USA	32.687528	-103.217028	POINT (-103.21703 32.68753)	New Mexico
2	ROW	Roswell Industrial Air Center	Roswell	NM	USA	33.301556	-104.530556	POINT (-104.53056 33.30156)	New Mexico
3	SAF	Santa Fe Municipal	Santa Fe	NM	USA	35.616778	-106.088139	POINT (-106.08814 35.61678)	New Mexico
4	AUS	Austin-Bergstrom International	Austin	TX	USA	30.194533	-97.669872	POINT (-97.66987 30.19453)	Texas

Spatial Joins

```
In [58]: ▶ us_states.plot()  
plt.show()
```



```
In [61]: ▶ us_states.head(15)
```

Out[61]:

	NAME_1	geometry
0	Alabama	MULTIPOLYGON (((-88.11320 30.22623, -88.11291 ...
1	Alaska	MULTIPOLYGON (((-141.31459 60.05416, -141.3125...
2	Arizona	POLYGON ((-110.53930 37.00423, -110.47991 37.0...
3	Arkansas	POLYGON ((-93.36903 36.49686, -93.36530 36.496...
4	California	MULTIPOLYGON (((-117.23285 32.77641, -117.2330...
5	Colorado	POLYGON ((-104.13991 41.00190, -104.05284 41.0...
6	Connecticut	MULTIPOLYGON (((-73.65778 40.98278, -73.65781 ...
7	Delaware	MULTIPOLYGON (((-75.09446 38.43233, -75.09415 ...
8	District of Columbia	POLYGON ((-77.00851 38.96956, -76.98849 38.954...
9	Florida	MULTIPOLYGON (((-81.96347 24.52542, -81.96347 ...
10	Georgia	MULTIPOLYGON (((-81.48306 30.77667, -81.48222 ...
11	Hawaii	MULTIPOLYGON (((-155.84998 20.26800, -155.8494...
12	Idaho	POLYGON ((-116.05072 49.00005, -116.05074 48.8...
13	Illinois	POLYGON ((-89.97443 42.50811, -89.95730 42.507...
14	Indiana	POLYGON ((-85.66060 41.76081, -85.57629 41.760...

```
In [62]: ▶ airport_us = airport_us[['AIRPORT', 'geometry']]
```

```
In [65]: ▶ airport_us
```

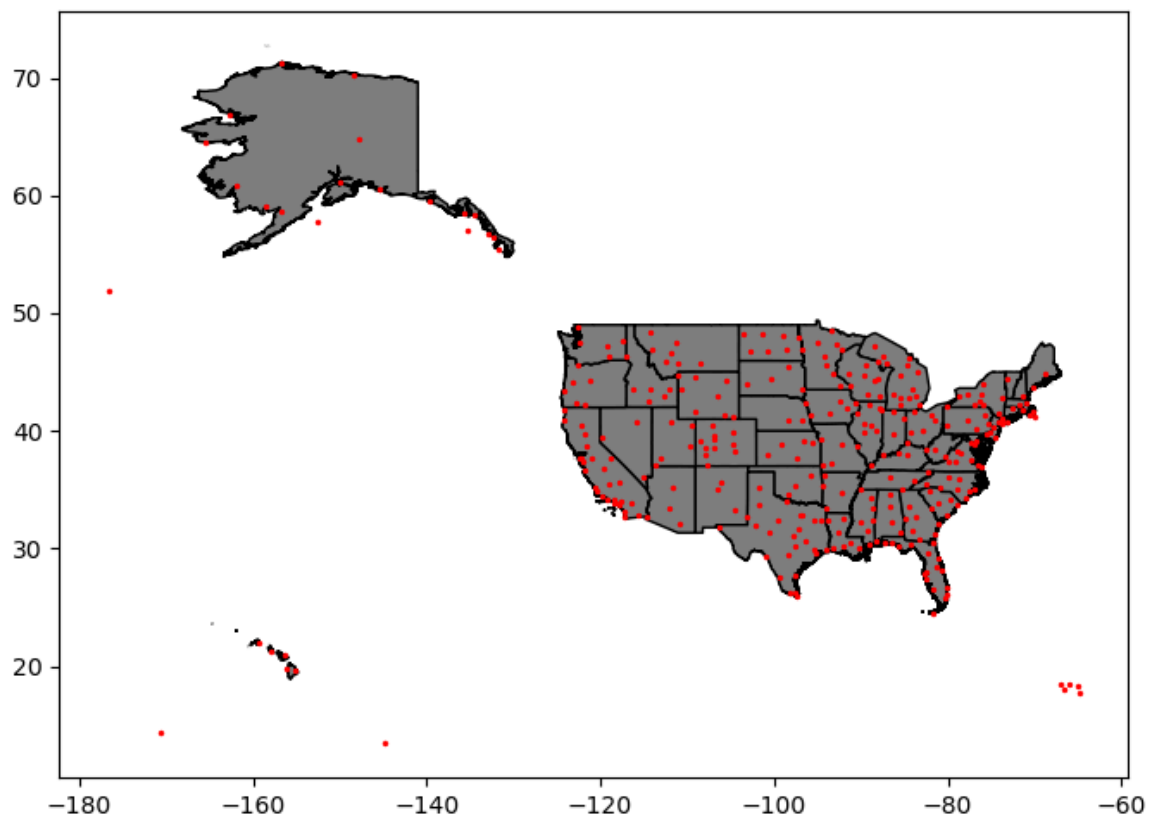
Out[65]:

	AIRPORT	geometry
0	Albuquerque International	POINT (-106.60919 35.04022)
1	Ted Stevens Anchorage International	POINT (-149.99619 61.17432)
2	William B Hartsfield-Atlanta Intl	POINT (-84.42694 33.64044)
3	Austin-Bergstrom International	POINT (-97.66987 30.19453)
4	Bradley International	POINT (-72.68323 41.93887)
...
335	Wrangell	POINT (-132.36982 56.48433)
336	Yellowstone	POINT (-111.11764 44.68840)
337	Northwest Arkansas Regional	POINT (-94.30681 36.28187)
338	Yakutat	POINT (-139.66023 59.50336)
339	Yuma MCAS-Yuma International	POINT (-114.60597 32.65658)

340 rows × 2 columns

```
In [70]: ▶ fig, ax = plt.subplots(figsize=(8, 8))
us_states.plot(ax=ax, color='gray', edgecolor='black')
airport_data.plot(ax=ax, markersize=2, color='red')
```

Out[70]: <Axes: >



```
In [71]: ▶ # Pefroming Special Join
```

```
In [74]: ▶ airport_us = gbd.sjoin(airport_us,us_states, how= 'inner', op= 'intersects')

C:\Users\DELL\anaconda3\Lib\site-packages\IPython\core\interactiveshell.py:3466: Future
Warning: The `op` parameter is deprecated and will be removed in a future release. Plea
se use the `predicate` parameter instead.
  if await self.run_code(code, result, async_=asy):
```

```
In [75]: ▶ airport_us
```

Out[75]:

	AIRPORT	geometry	index_right	NAME_1
0	Albuquerque International	POINT (-106.60919 35.04022)	31	New Mexico
196	Lea County Regional	POINT (-103.21703 32.68753)	31	New Mexico
295	Roswell Industrial Air Center	POINT (-104.53056 33.30156)	31	New Mexico
297	Santa Fe Municipal	POINT (-106.08814 35.61678)	31	New Mexico
1	Ted Stevens Anchorage International	POINT (-149.99619 61.17432)	1	Alaska
...
242	Key	POINT (-88.75121 32.33313)	24	Mississippi
274	Hattiesburg-Laurel Regional	POINT (-89.33706 31.46715)	24	Mississippi
326	Tupelo Municipal	POINT (-88.76990 34.26811)	24	Mississippi
205	New Castle County	POINT (-75.60653 39.67872)	7	Delaware
247	Manchester	POINT (-71.43706 42.93452)	29	New Hampshire

326 rows × 4 columns

```
In [ ]: ▶
```