Learning_Pandas_Part_104_SQL_equivalent_in_Pandas

February 2, 2022

0.0.1 Prepared by Abhishek Kumar

0.0.2 https://www.linkedin.com/in/abhishekkumar-0311/

[16]:
```
# To get multiple outputs in the same cell

from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"

%matplotlib inline
```

[17]:
```
import sqlite3 as sql
```

[18]:
```
import pandas as pd
import numpy as np
import sqlite3 as sql
# conn = sql.connect('default.db')
import matplotlib.pyplot as plt
```

[39]:
```

```

I got this data at http://ourairports.com/data/.

[21]:
```
airport = pd.read_csv('./data/airports.csv')
run = pd.read_csv('./data/runways.csv')
```

[22]:
```
airport.shape
run.shape
```

[22]: (69063, 18)

[22]: (42895, 20)

[57]:
```
# df_return = sqldb(airport,'airport',query,conn=, direct_run=)
def sqldb(sql_query, conn = sql.connect('default.db'), direct_run = 1, df = df,␣
 ↪sql_tbl = sql_tbl ):
    ''' df -> pandas dataframe
        sql_tbl -> equivalent table in Db
```

```
        sql_query -> query to be executed in sql env
        conn -> connection to database: deafult db is set to default.db
        direct_run -> indicates whether the query needs to be executed on␣
↪existing table or needs to be re-created '''
    if direct_run == 0:
        cursor = conn.cursor()
        drop_query = "DROP TABLE IF EXISTS " + sql_tbl
        cursor.execute(drop_query)
        df.to_sql(sql_tbl,conn)
    dfinit = pd.read_sql(sql_query,conn)
    return dfinit
```

[59]:
```
conn = sql.connect('default.db')
direct_run = 1
df_arg = airport
sql_tbl = 'airport'
query = 'select * from airport limit 10'

df_return = sqldb(query,conn,direct_run=1)
df_return
```

[59]:

	level_0	index	id	ident	type	name	latitude_deg	longitude_deg	elevation_ft	continent	iso_country	iso_region	municipality
0	0	0	6523	00A	heliport	Total Rf Heliport	40.070801	-74.933601	11.0	None	US	US-PA	Bensalem
1	1	1	323361	00AA	small_airport	Aero B Ranch Airport	38.704022	-101.473911	3435.0	None	US	US-KS	Leoti
2	2	2	6524	00AK	small_airport	Lowell Field	59.947733	-151.692524	450.0	None	US	US-AK	Anchor Point
3	3	3	6525	00AL	small_airport	Epps Airpark	34.864799	-86.770302	820.0	None	US	US-AL	Harvest
4	4	4	6526	00AR	closed	Newport Hospital & Clinic Heliport	35.608700	-91.254898	237.0	None	US	US-AR	Newport
5	5	5	322127	00AS	small_airport	Fulton Airport	34.942803	-97.818019	1100.0	None	US	US-OK	Alex
6	6	6	6527	00AZ	small_airport	Cordes Airport	34.305599	-112.165001	3810.0	None	US	US-AZ	Cordes
7	7	7	6528	00CA	small_airport	Goldstone (GTS) Airport	35.354740	-116.885329	3038.0	None	US	US-CA	Barstow
8	8	8	324424	00CL	small_airport	Williams Ag Airport	39.427188	-121.763427	87.0	None	US	US-CA	Biggs
9	9	9	322658	00CN	heliport	Kitchen Creek Helibase Heliport	32.727374	-116.459742	3350.0	None	US	US-CA	Pine Valley

```
1        3435.0      None          US      US-KS         Leoti
2         450.0      None          US      US-AK  Anchor Point
3         820.0      None          US      US-AL       Harvest
4         237.0      None          US      US-AR       Newport
5        1100.0      None          US      US-OK          Alex
6        3810.0      None          US      US-AZ        Cordes
7        3038.0      None          US      US-CA       Barstow
8          87.0      None          US      US-CA         Biggs
9        3350.0      None          US      US-CA   Pine Valley

  scheduled_service gps_code iata_code local_code home_link wikipedia_link  \
0                no      00A      None        00A      None           None
1                no     00AA      None       00AA      None           None
2                no     00AK      None       00AK      None           None
3                no     00AL      None       00AL      None           None
4                no     None      None       None      None           None
5                no     00AS      None       00AS      None           None
6                no     00AZ      None       00AZ      None           None
7                no     00CA      None       00CA      None           None
8                no     00CL      None       00CL      None           None
9                no     00CN      None       00CN      None           None

  keywords
0     None
1     None
2     None
3     None
4     00AR
5     None
6     None
7     None
8     None
9     None
```

[54]:
```
%reset?
```

[13]:
```
airport.to_sql('airport',conn)
query = 'select * from airport limit 10'
dfinit = pd.read_sql(query,conn)
dfinit
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-13-faac85060a46> in <module>
----> 1 airport.to_sql('airport',conn)
      2 query = 'select * from airport limit 10'
      3 dfinit = pd.read_sql(query,conn)
```

```
      4 dfinit

NameError: name 'airport' is not defined
```

[41]:
```
ap = airport.copy()
ap.head(3)
```

[41]:
```
       id ident           type                   name  latitude_deg  \
0    6523   00A       heliport     Total Rf Heliport     40.070801
1  323361  00AA  small_airport  Aero B Ranch Airport     38.704022
2    6524  00AK  small_airport          Lowell Field     59.947733

   longitude_deg  elevation_ft continent iso_country iso_region  municipality  \
0     -74.933601          11.0       NaN          US      US-PA      Bensalem
1    -101.473911        3435.0       NaN          US      US-KS         Leoti
2    -151.692524         450.0       NaN          US      US-AK  Anchor Point

  scheduled_service gps_code iata_code local_code home_link wikipedia_link  \
0                no      00A       NaN        00A       NaN            NaN
1                no     00AA       NaN       00AA       NaN            NaN
2                no     00AK       NaN       00AK       NaN            NaN

  keywords
0      NaN
1      NaN
2      NaN
```

[ ]:

[ ]:

[ ]:

[ ]:

1 Sample SQL

- SELECT… FROM… WHERE…
- GROUP BY… HAVING…
- ORDER BY…
- LIMIT… OFFSET…

1.1 SELECT... FROM...

[56]:
```
# SELECT… < *, Few> FROM…

ap                                  # SELECT *
ap.head(3)                          # LIMIT
ap[['ident','type','name']]         # SELECT Few
ap.type.unique()                    # SELECT DISTINCT
list(ap.type.unique())              # SELECT DISTINCT -> to get distinct values␣
 ↪as a list
ap.type.nunique()                   # SELECT DISTINCT COUNT
len(ap.type.unique())               # SELECT DISTINCT COUNT
```

[56]:

	id	ident	type	name	\
0	6523	00A	heliport	Total Rf Heliport	
1	323361	00AA	small_airport	Aero B Ranch Airport	
2	6524	00AK	small_airport	Lowell Field	
3	6525	00AL	small_airport	Epps Airpark	
4	6526	00AR	closed	Newport Hospital & Clinic Heliport	
…	…	…	…	…	
69058	46378	ZZ-0001	heliport	Sealand Helipad	
69059	307326	ZZ-0002	small_airport	Glorioso Islands Airstrip	
69060	346788	ZZ-0003	small_airport	Fainting Goat Airport	
69061	342102	ZZZW	closed	Scandium City Heliport	
69062	313629	ZZZZ	small_airport	Satsuma Iōjima Airport	

	latitude_deg	longitude_deg	elevation_ft	continent	iso_country	\
0	40.070801	-74.933601	11.0	NaN	US	
1	38.704022	-101.473911	3435.0	NaN	US	
2	59.947733	-151.692524	450.0	NaN	US	
3	34.864799	-86.770302	820.0	NaN	US	
4	35.608700	-91.254898	237.0	NaN	US	
…	…	…	…	…	…	
69058	51.894444	1.482500	40.0	EU	GB	
69059	-11.584278	47.296389	11.0	AF	TF	
69060	32.110587	-97.356312	690.0	NaN	US	
69061	69.355287	-138.939310	4.0	NaN	CA	
69062	30.784722	130.270556	338.0	AS	JP	

	iso_region	municipality	scheduled_service	gps_code	iata_code	\
0	US-PA	Bensalem	no	00A	NaN	
1	US-KS	Leoti	no	00AA	NaN	
2	US-AK	Anchor Point	no	00AK	NaN	
3	US-AL	Harvest	no	00AL	NaN	
4	US-AR	Newport	no	NaN	NaN	
…	…	…	…	…	…	
69058	GB-ENG	Sealand	no	NaN	NaN	

```
69059     TF-U-A     Grande Glorieuse                no      NaN       NaN
69060      US-TX                 Blum                no     87TX       NaN
69061      CA-YT  (Old) Scandium City                no     ZZZW       ZYW
69062      JP-46              Mishima                no     RJX7       NaN

      local_code                    home_link  \
0            00A                          NaN
1           00AA                          NaN
2           00AK                          NaN
3           00AL                          NaN
4            NaN                          NaN
…            …                            …
69058        NaN  http://www.sealandgov.org/
69059        NaN                          NaN
69060       87TX                          NaN
69061       YK96                          NaN
69062       RJX7                          NaN

                                        wikipedia_link              keywords
0                                                  NaN                   NaN
1                                                  NaN                   NaN
2                                                  NaN                   NaN
3                                                  NaN                   NaN
4                                                  NaN                 00AR
…                                                  …                     …
69058  https://en.wikipedia.org/wiki/Principality_of_…  Roughs Tower Helipad
69059                                              NaN                   NaN
69060                                              NaN                   NaN
69061                                              NaN                   NaN
69062  http://wikimapia.org/6705190/Satsuma-Iwo-jima-…  SATSUMA,IWOJIMA,RJX7

[69063 rows x 18 columns]
```

[56]:
```
       id ident          type                   name  latitude_deg  \
0    6523   00A      heliport     Total Rf Heliport     40.070801
1  323361  00AA  small_airport  Aero B Ranch Airport     38.704022
2    6524  00AK  small_airport          Lowell Field     59.947733

   longitude_deg  elevation_ft continent iso_country iso_region  municipality  \
0     -74.933601          11.0       NaN          US      US-PA      Bensalem
1    -101.473911        3435.0       NaN          US      US-KS         Leoti
2    -151.692524         450.0       NaN          US      US-AK  Anchor Point

  scheduled_service gps_code iata_code local_code home_link wikipedia_link  \
0                no      00A       NaN        00A       NaN            NaN
1                no     00AA       NaN       00AA       NaN            NaN
2                no     00AK       NaN       00AK       NaN            NaN
```

```
  keywords
0      NaN
1      NaN
2      NaN
```

[56]:
```
          ident           type                                name
0           00A       heliport                   Total Rf Heliport
1          00AA  small_airport                Aero B Ranch Airport
2          00AK  small_airport                        Lowell Field
3          00AL  small_airport                        Epps Airpark
4          00AR         closed  Newport Hospital & Clinic Heliport
…             …              …                                   …
69058   ZZ-0001       heliport                     Sealand Helipad
69059   ZZ-0002  small_airport           Glorioso Islands Airstrip
69060   ZZ-0003  small_airport               Fainting Goat Airport
69061      ZZZW         closed              Scandium City Heliport
69062      ZZZZ  small_airport              Satsuma Iōjima Airport

[69063 rows x 3 columns]
```

[56]:
```
array(['heliport', 'small_airport', 'closed', 'seaplane_base',
       'balloonport', 'medium_airport', 'large_airport'], dtype=object)
```

[56]:
```
['heliport',
 'small_airport',
 'closed',
 'seaplane_base',
 'balloonport',
 'medium_airport',
 'large_airport']
```

[56]:
```
7
```

[56]:
```
7
```

1.2 SELECT… FROM… WHERE… ORDER BY… and LIMIT…

[61]:
```
# SELECT… FROM… WHERE… ORDER BY… and LIMIT…

ap[['ident','type','name']][ap.type == 'small_airport']
 ↪# SELECT Few & WHERE
ap[['ident','type','name']][ap.type == 'small_airport'].head(3)
 ↪# SELECT Few & WHERE and LIMIT
ap[ap.type == 'small_airport'].head(3)
 ↪# SELECT ALL & WHERE (==) and LIMIT
```

```
ap[ap.type.isin(['small_airport','heliport'])][['ident','type','name']].head(3)␣
 ↪# SELECT Few & WHERE (.isin()) and LIMIT
ap[ap.type == 'small_airport'].head(3).sort_values('name', ascending = False)  ␣
 ↪# SELECT ALL & WHERE (==) and LIMIT and ORDER BY
ap[ap.type == 'small_airport'].sort_values('name', ascending = False).head(3)  ␣
 ↪# SELECT ALL & WHERE (==) and LIMIT and ORDER BY
```

[61]:

	ident	type	name
1	00AA	small_airport	Aero B Ranch Airport
2	00AK	small_airport	Lowell Field
3	00AL	small_airport	Epps Airpark
5	00AS	small_airport	Fulton Airport
6	00AZ	small_airport	Cordes Airport
…	…	…	…
69045	ZYLY	small_airport	Liaoyang Airbase
69050	ZYTH	small_airport	Tahe Airport
69059	ZZ-0002	small_airport	Glorioso Islands Airstrip
69060	ZZ-0003	small_airport	Fainting Goat Airport
69062	ZZZZ	small_airport	Satsuma Iōjima Airport

[37666 rows x 3 columns]

[61]:

	ident	type	name
1	00AA	small_airport	Aero B Ranch Airport
2	00AK	small_airport	Lowell Field
3	00AL	small_airport	Epps Airpark

[61]:

	id	ident	type	name	latitude_deg	\
1	323361	00AA	small_airport	Aero B Ranch Airport	38.704022	
2	6524	00AK	small_airport	Lowell Field	59.947733	
3	6525	00AL	small_airport	Epps Airpark	34.864799	

	longitude_deg	elevation_ft	continent	iso_country	iso_region	municipality	\
1	-101.473911	3435.0	NaN	US	US-KS	Leoti	
2	-151.692524	450.0	NaN	US	US-AK	Anchor Point	
3	-86.770302	820.0	NaN	US	US-AL	Harvest	

	scheduled_service	gps_code	iata_code	local_code	home_link	wikipedia_link	\
1	no	00AA	NaN	00AA	NaN	NaN	
2	no	00AK	NaN	00AK	NaN	NaN	
3	no	00AL	NaN	00AL	NaN	NaN	

	keywords
1	NaN
2	NaN
3	NaN

[61]:
```
  ident           type                  name
0   00A       heliport     Total Rf Heliport
1  00AA  small_airport  Aero B Ranch Airport
2  00AK  small_airport          Lowell Field
```

[61]:
```
        id ident           type                  name  latitude_deg  \
2     6524  00AK  small_airport          Lowell Field     59.947733
3     6525  00AL  small_airport          Epps Airpark     34.864799
1   323361  00AA  small_airport  Aero B Ranch Airport     38.704022

    longitude_deg  elevation_ft continent iso_country iso_region  municipality  \
2    -151.692524         450.0       NaN          US      US-AK  Anchor Point
3     -86.770302         820.0       NaN          US      US-AL       Harvest
1    -101.473911        3435.0       NaN          US      US-KS         Leoti

  scheduled_service gps_code iata_code local_code home_link wikipedia_link  \
2                no     00AK       NaN       00AK       NaN            NaN
3                no     00AL       NaN       00AL       NaN            NaN
1                no     00AA       NaN       00AA       NaN            NaN

  keywords
2      NaN
3      NaN
1      NaN
```

[61]:
```
            id    ident           type              name  latitude_deg  \
38915  325859  LV-0026  small_airport  Žocenes lidlauks     57.532464
52525  316302  SK-0011  small_airport        Želiezovce     48.047222
19043  321024  CZ-0131  small_airport    Želeč Airstrip     49.330555

       longitude_deg  elevation_ft continent iso_country iso_region  \
38915      22.699390           NaN        EU          LV     LV-U-A
52525      18.596111           NaN        EU          SK      SK-NJ
19043      14.658889        1513.0        EU          CZ      CZ-JC

      municipality scheduled_service gps_code iata_code local_code home_link  \
38915       Žocene                no      NaN       NaN        NaN       NaN
52525          NaN                no      NaN       NaN        NaN       NaN
19043          NaN                no      NaN       NaN        NaN       NaN

      wikipedia_link keywords
38915            NaN      NaN
52525            NaN      NaN
19043            NaN      NaN
```

1.3 SELECT… FROM… WHERE… GROUP BY… HAVING… ORDER BY… and LIMIT…

[63]:
```
run.head(3)
```

[63]:
```
       id  airport_ref airport_ident  length_ft  width_ft surface  lighted  \
0  269408         6523           00A       80.0      80.0  ASPH-G        1
1  255155         6524          00AK     2500.0      70.0    GRVL        0
2  254165         6525          00AL     2300.0     200.0    TURF        0

   closed le_ident  le_latitude_deg  le_longitude_deg  le_elevation_ft  \
0       0       H1              NaN               NaN              NaN
1       0        N              NaN               NaN              NaN
2       0       01              NaN               NaN              NaN

   le_heading_degT  le_displaced_threshold_ft he_ident  he_latitude_deg  \
0              NaN                        NaN      NaN              NaN
1              NaN                        NaN        S              NaN
2              NaN                        NaN       19              NaN

   he_longitude_deg  he_elevation_ft  he_heading_degT  \
0               NaN              NaN              NaN
1               NaN              NaN              NaN
2               NaN              NaN              NaN

   he_displaced_threshold_ft
0                        NaN
1                        NaN
2                        NaN
```

[68]:
```
run.airport_ref.value_counts()
```

[68]:
```
300162    10
3754      10
18256      9
8413       8
24720      8
          ..
302215     1
322697     1
335223     1
9360       1
2049       1
Name: airport_ref, Length: 36275, dtype: int64
```

[70]:
```
t = run.groupby('airport_ref')
```

[78]:
```
t.get_group('3754')
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
<ipython-input-78-fd898ded808b> in <module>
----> 1 t.get_group('3754')

~\anaconda3\lib\site-packages\pandas\core\groupby\groupby.py in get_group(self,
 ↪name, obj)
    808         inds = self._get_index(name)
    809         if not len(inds):
--> 810             raise KeyError(name)
    811
    812         return obj._take_with_is_copy(inds, axis=self.axis)

KeyError: '3754'
```

[79]:
```
t.first()
```

[79]:

airport_ref	id	airport_ident	length_ft	width_ft	surface	lighted	closed	le_ident	le_latitude_deg	le_longitude_deg
2	232775	OM11	7080.0	200.0	ASP	0	0	13	24.52520	54.9720
3	233754	AGGH	7218.0	148.0	ASP	1	0	06	-9.43174	160.0450
4	233755	AGGM	4593.0	98.0	ASP	0	0	07	-8.32918	157.2570
5	232874	AL-LA10	3855.0	228.0	GRS	0	0	14	40.09150	20.1489
7	232877	UD21	3281.0	250.0	CON	0	0	05	40.29630	44.5561
…	…	…	…	…	…	…	…	…	…	…
349923	349954	SDK9	2625.0	59.0	GRS	0	0	10	NaN	NaN
349941	349953	SDL7	2067.0	59.0	ASP	0	0	03	NaN	NaN
349943	349951	SDL8	3281.0	59.0	GRS	0	0	11	NaN	NaN
349944	349952	SDL9	2625.0	59.0	GRS	0	0	06	NaN	NaN
349949	349950	SDM9	4921.0	66.0	ASP	0	0	10	NaN	NaN

```
             le_elevation_ft  le_heading_degT  le_displaced_threshold_ft  \
airport_ref
2                        NaN            130.9                        NaN
3                       28.0             68.0                        NaN
4                        7.0             78.9                        NaN
5                        NaN            142.0                        NaN
7                        NaN             60.3                        NaN
…                        …               …                          …
349923                   NaN              NaN                        NaN
349941                   NaN              NaN                        NaN
349943                   NaN              NaN                        NaN
349944                   NaN              NaN                        NaN
349949                   NaN              NaN                        NaN

            he_ident  he_latitude_deg  he_longitude_deg  he_elevation_ft  \
airport_ref
2                 31         24.51250           54.9881              NaN
3                 24         -9.42426          160.0640             14.0
4                 25         -8.32676          157.2690              9.0
5                 32         40.08310           20.1573              NaN
7                 23         40.30070           44.5663              NaN
…                 …               …                 …                …
349923            28              NaN               NaN              NaN
349941            21              NaN               NaN              NaN
349943            29              NaN               NaN              NaN
349944            24              NaN               NaN              NaN
349949            28              NaN               NaN              NaN

             he_heading_degT  he_displaced_threshold_ft
airport_ref
2                      310.9                        NaN
3                      248.0                        NaN
4                      258.9                        NaN
5                      322.0                        NaN
7                      240.3                        NaN
…                      …                            …
349923                   NaN                        NaN
349941                   NaN                        NaN
349943                   NaN                        NaN
349944                   NaN                        NaN
349949                   NaN                        NaN

[36275 rows x 19 columns]
```

[80]:
```
t.head(1)
```

[80]:

	id	airport_ref	airport_ident	length_ft	width_ft	surface	\
0	269408	6523	00A	80.0	80.0	ASPH-G	
1	255155	6524	00AK	2500.0	70.0	GRVL	
2	254165	6525	00AL	2300.0	200.0	TURF	
3	270932	6526	00AR	40.0	40.0	GRASS	
4	322128	322127	00AS	1450.0	60.0	Turf	
…	…	…	…	…	…	…	
42890	235188	27242	ZYTL	10827.0	148.0	CON	
42891	235186	27243	ZYTX	10499.0	148.0	Asphalt	
42892	235169	27244	ZYYJ	8530.0	148.0	CON	
42893	346789	346788	ZZ-0003	1800.0	15.0	Turf	
42894	313663	313629	ZZZZ	1713.0	82.0	concrete	

	lighted	closed	le_ident	le_latitude_deg	le_longitude_deg	\
0	1	0	H1	NaN	NaN	
1	0	0	N	NaN	NaN	
2	0	0	01	NaN	NaN	
3	0	0	H1	NaN	NaN	
4	0	0	1	NaN	NaN	
…	…	…	…	…	…	
42890	1	0	10	38.9671	121.520	
42891	1	0	06	41.6304	123.469	
42892	1	0	09	42.8811	129.436	
42893	0	0	15	NaN	NaN	
42894	0	0	18	30.7835	130.273	

	le_elevation_ft	le_heading_degT	le_displaced_threshold_ft	he_ident	\
0	NaN	NaN	NaN	NaN	
1	NaN	NaN	NaN	S	
2	NaN	NaN	NaN	19	
3	NaN	NaN	NaN	H1	
4	NaN	NaN	NaN	19	
…	…	…	…	…	
42890	105.0	95.6	650.0	28	
42891	171.0	49.0	NaN	24	
42892	623.0	81.7	NaN	27	
42893	NaN	NaN	NaN	33	
42894	NaN	NaN	NaN	36	

	he_latitude_deg	he_longitude_deg	he_elevation_ft	he_heading_degT	\
0	NaN	NaN	NaN	NaN	
1	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	
…	…	…	…	…	
42890	38.9642	121.558	85.0	275.6	

```
42891          41.6493          123.498            197.0            229.0
42892          42.8845          129.467            597.0            261.7
42893              NaN              NaN              NaN              NaN
42894          30.7781          130.273              NaN              NaN

       he_displaced_threshold_ft
0                            NaN
1                            NaN
2                            NaN
3                            NaN
4                            NaN
…                              …
42890                      320.0
42891                        NaN
42892                        NaN
42893                        NaN
42894                        NaN

[36275 rows x 20 columns]
```

[82]:
```
t.ngroup()
```

[82]:
```
0         6017
1         6018
2         6019
3         6020
4        34495
         …
42890    25474
42891    25475
42892    25476
42893    36006
42894    33980
Length: 42895, dtype: int64
```

[83]:
```
t.nth(2)
```

[83]:
```
                id airport_ident  length_ft  width_ft surface  lighted  \
airport_ref
24          232919          SA14     1968.0      98.0     GRE        0
41          249970          SA34     5739.0     131.0     GRE        0
125         236459          BIRK     4035.0     148.0     ASP        1
647         235035          CAU4     3200.0     180.0     GVL        0
773         234648          CCQ3     5000.0     150.0     ASP        0
…              …             …          …         …       …      …
330312      330315       ZA-0161      900.0       NaN   Grass        0
330820      330822          ZBAD    12467.0     197.0     CON        1
```

```
334256       334259       US-1262     3000.0     100.0   Water        0
345603       345606          40NK       30.0      30.0     CON        0
347732       347765          96MT     1357.0      50.0    Turf        0

             closed le_ident  le_latitude_deg  le_longitude_deg  \
airport_ref
24                0        9         -38.2250          -57.8799
41                0       13         -43.2268          -65.3355
125               0       13          64.1325          -21.9554
647               0       15          54.0453         -124.0210
773               0       16          45.4253          -63.4703
…              …       …              …                …
330312            0       13              NaN               NaN
330820            0      17L          39.5179          116.3960
334256            0       8W              NaN               NaN
345603            0       H3              NaN               NaN
347732            0        7              NaN               NaN

             le_elevation_ft  le_heading_degT  le_displaced_threshold_ft  \
airport_ref
24                       NaN            96.0                        NaN
41                       NaN           131.0                        NaN
125                     17.0           117.0                        NaN
647                      NaN           172.0                      100.0
773                    130.0           138.8                      500.0
…                        …              …                          …
330312                5001.0            13.0                     5001.0
330820                  76.0             NaN                        NaN
334256                   0.0             NaN                        NaN
345603                   NaN             NaN                        NaN
347732                   NaN             NaN                        NaN

            he_ident  he_latitude_deg  he_longitude_deg  he_elevation_ft  \
airport_ref
24                27         -38.2264          -57.8632              NaN
41                31         -43.2371          -65.3193              NaN
125               31          64.1276          -21.9327             38.0
647               33          54.0366         -124.0190              NaN
773               34          45.4150          -63.4575            135.0
…               …              …                …                …
330312            31             NaN               NaN           5001.0
330820           35R          39.4839          116.4010             76.0
334256           26W             NaN               NaN              0.0
345603           NaN             NaN               NaN              NaN
347732            25             NaN               NaN              NaN

             he_heading_degT  he_displaced_threshold_ft
```

```
airport_ref
24                    276.0                        NaN
41                    311.0                        NaN
125                   297.0                      213.0
647                   352.0                        NaN
773                   318.8                        NaN
…                     …                            …
330312                 31.0                     5001.0
330820                  NaN                        NaN
334256                  NaN                        NaN
345603                  NaN                        NaN
347732                  NaN                        NaN

[1089 rows x 19 columns]
```

[84]:
```
t.size()
```

[84]:
```
airport_ref
2         1
3         1
4         1
5         1
7         2
         ..
349923    1
349941    1
349943    1
349944    1
349949    1
Length: 36275, dtype: int64
```

[86]:
```
t.all()
```

[86]:
```
                id  airport_ident  length_ft  width_ft  surface  lighted  \
airport_ref
2             True           True       True      True     True    False
3             True           True       True      True     True     True
4             True           True       True      True     True    False
5             True           True       True      True     True    False
7             True           True       True      True     True    False
…             …              …          …         …        …       …
349923        True           True       True      True     True    False
349941        True           True       True      True     True    False
349943        True           True       True      True     True    False
349944        True           True       True      True     True    False
349949        True           True       True      True     True    False
```

```
             closed  le_ident  le_latitude_deg  le_longitude_deg  \
airport_ref
2             False      True             True              True
3             False      True             True              True
4             False      True             True              True
5             False      True             True              True
7             False      True             True              True
…               …       …                …                 …
349923        False      True             True              True
349941        False      True             True              True
349943        False      True             True              True
349944        False      True             True              True
349949        False      True             True              True

             le_elevation_ft  le_heading_degT  le_displaced_threshold_ft  \
airport_ref
2                       True             True                       True
3                       True             True                       True
4                       True             True                       True
5                       True             True                       True
7                       True             True                       True
…                       …                …                          …
349923                  True             True                       True
349941                  True             True                       True
349943                  True             True                       True
349944                  True             True                       True
349949                  True             True                       True

             he_ident  he_latitude_deg  he_longitude_deg  he_elevation_ft  \
airport_ref
2                True             True              True             True
3                True             True              True             True
4                True             True              True             True
5                True             True              True             True
7                True             True              True             True
…                …                …                 …                …
349923           True             True              True             True
349941           True             True              True             True
349943           True             True              True             True
349944           True             True              True             True
349949           True             True              True             True

             he_heading_degT  he_displaced_threshold_ft
airport_ref
2                       True                       True
3                       True                       True
4                       True                       True
```

```
5                      True                      True
7                      True                      True
…                      …                         …
349923                 True                      True
349941                 True                      True
349943                 True                      True
349944                 True                      True
349949                 True                      True

[36275 rows x 19 columns]
```

[87]:
```
t.nunique()
```

[87]:
```
             id  airport_ident  length_ft  width_ft  surface  lighted  closed  \
airport_ref
2             1              1          1         1        1        1       1
3             1              1          1         1        1        1       1
4             1              1          1         1        1        1       1
5             1              1          1         1        1        1       1
7             2              1          2         1        1        1       1
…           ..             …          …         …        …        …       …
349923        1              1          1         1        1        1       1
349941        1              1          1         1        1        1       1
349943        1              1          1         1        1        1       1
349944        1              1          1         1        1        1       1
349949        1              1          1         1        1        1       1

             le_ident  le_latitude_deg  le_longitude_deg  le_elevation_ft  \
airport_ref
2                   1                1                 1                0
3                   1                1                 1                1
4                   1                1                 1                1
5                   1                1                 1                0
7                   2                2                 2                0
…                 …                …                 …                …
349923              1                0                 0                0
349941              1                0                 0                0
349943              1                0                 0                0
349944              1                0                 0                0
349949              1                0                 0                0

             le_heading_degT  le_displaced_threshold_ft  he_ident  \
airport_ref
2                          1                          0         1
3                          1                          0         1
4                          1                          0         1
5                          1                          0         1
```

```
7                          2                          0          2
…                        …                          …        …
349923                     0                          0          1
349941                     0                          0          1
349943                     0                          0          1
349944                     0                          0          1
349949                     0                          0          1

             he_latitude_deg  he_longitude_deg  he_elevation_ft  \
airport_ref
2                          1                 1                0
3                          1                 1                1
4                          1                 1                1
5                          1                 1                0
7                          2                 2                0
…                        …                 …                …
349923                     0                 0                0
349941                     0                 0                0
349943                     0                 0                0
349944                     0                 0                0
349949                     0                 0                0

             he_heading_degT  he_displaced_threshold_ft
airport_ref
2                          1                          0
3                          1                          0
4                          1                          0
5                          1                          0
7                          2                          0
…                        …                          …
349923                     0                          0
349941                     0                          0
349943                     0                          0
349944                     0                          0
349949                     0                          0

[36275 rows x 19 columns]
```

[88]:
```
t.describe()
```

```
---------------------------------------------------------------------------
KeyboardInterrupt                         Traceback (most recent call last)
<ipython-input-88-aceb2410b4df> in <module>
----> 1 t.describe()

~\anaconda3\lib\site-packages\pandas\core\groupby\groupby.py in describe(self,␣
 ↪**kwargs)
```

```
   1627     def describe(self, **kwargs):
   1628         with _group_selection_context(self):
-> 1629             result = self.apply(lambda x: x.describe(**kwargs))
   1630             if self.axis == 1:
   1631                 return result.T

~\anaconda3\lib\site-packages\pandas\core\groupby\groupby.py in apply(self,␣
↪func, *args, **kwargs)
    857         with option_context("mode.chained_assignment", None):
    858             try:
--> 859                 result = self._python_apply_general(f, self.
↪_selected_obj)
    860             except TypeError:
    861                 # gh-20949

~\anaconda3\lib\site-packages\pandas\core\groupby\groupby.py in␣
↪_python_apply_general(self, f, data)
    890             data after applying f
    891         """
--> 892         keys, values, mutated = self.grouper.apply(f, data, self.axis)
    893
    894         return self._wrap_applied_output(

~\anaconda3\lib\site-packages\pandas\core\groupby\ops.py in apply(self, f, data␣
↪axis)
    177         ):
    178             try:
--> 179                 result_values, mutated = splitter.fast_apply(f, sdata,␣
↪group_keys)
    180
    181             except libreduction.InvalidApply as err:

~\anaconda3\lib\site-packages\pandas\core\groupby\ops.py in fast_apply(self, f,
↪sdata, names)
    963         # must return keys::list, values::list, mutated::bool
    964         starts, ends = lib.generate_slices(self.slabels, self.ngroups)
--> 965         return libreduction.apply_frame_axis0(sdata, f, names, starts,␣
↪ends)
    966
    967     def _chop(self, sdata: DataFrame, slice_obj: slice) -> DataFrame:

pandas\_libs\reduction.pyx in pandas._libs.reduction.apply_frame_axis0()

~\anaconda3\lib\site-packages\pandas\core\groupby\groupby.py in <lambda>(x)
   1627     def describe(self, **kwargs):
   1628         with _group_selection_context(self):
-> 1629             result = self.apply(lambda x: x.describe(**kwargs))
   1630             if self.axis == 1:
```

```
   1631                 return result.T

~\anaconda3\lib\site-packages\pandas\core\generic.py in describe(self,␣
 ↪percentiles, include, exclude, datetime_is_numeric)
  10096             data = self.select_dtypes(include=include, exclude=exclude)
  10097
> 10098         ldesc = [describe_1d(s) for _, s in data.items()]
  10099         # set a convenient order for rows
  10100         names: List[Label] = []

~\anaconda3\lib\site-packages\pandas\core\generic.py in <listcomp>(.0)
  10096             data = self.select_dtypes(include=include, exclude=exclude)
  10097
> 10098         ldesc = [describe_1d(s) for _, s in data.items()]
  10099         # set a convenient order for rows
  10100         names: List[Label] = []

~\anaconda3\lib\site-packages\pandas\core\generic.py in describe_1d(data)
  10070                 return describe_categorical_1d(data)
  10071             elif is_numeric_dtype(data):
> 10072                 return describe_numeric_1d(data)
  10073             elif is_datetime64_any_dtype(data.dtype) and␣
 ↪datetime_is_numeric:
  10074                 return describe_timestamp_1d(data)

~\anaconda3\lib\site-packages\pandas\core\generic.py in␣
 ↪describe_numeric_1d(series)
  10002             d = (
  10003                 [series.count(), series.mean(), series.std(), series.
 ↪min()]
> 10004                 + series.quantile(percentiles).tolist()
  10005                 + [series.max()]
  10006             )

~\anaconda3\lib\site-packages\pandas\core\series.py in quantile(self, q,␣
 ↪interpolation)
   2259         df = self.to_frame()
   2260
-> 2261         result = df.quantile(q=q, interpolation=interpolation,␣
 ↪numeric_only=False)
   2262         if result.ndim == 2:
   2263             result = result.iloc[:, 0]

~\anaconda3\lib\site-packages\pandas\core\frame.py in quantile(self, q, axis,␣
 ↪numeric_only, interpolation)
   9057             return self._constructor_sliced([], index=cols, name=q,␣
 ↪dtype=np.float64)
   9058
```

```
-> 9059         result = data._mgr.quantile(
   9060             qs=q, axis=1, interpolation=interpolation,␣
 ↪transposed=is_transposed
   9061         )

~\anaconda3\lib\site-packages\pandas\core\internals\managers.py in␣
 ↪quantile(self, axis, transposed, interpolation, qs, numeric_only)
    455         axes, blocks = [], []
    456         for b in self.blocks:
--> 457             block = b.quantile(axis=axis, qs=qs,␣
 ↪interpolation=interpolation)
    458
    459             axe = get_axe(b, qs, axes=self.axes)

~\anaconda3\lib\site-packages\pandas\core\internals\blocks.py in quantile(self,
 ↪qs, interpolation, axis)
   1448             # asarray needed for Sparse, see GH#24600
   1449             mask = np.asarray(isna(values))
-> 1450             result = nanpercentile(
   1451                 values,
   1452                 np.array(qs) * 100,

~\anaconda3\lib\site-packages\pandas\core\nanops.py in nanpercentile(values, q,
 ↪axis, na_value, mask, ndim, interpolation)
   1556             return result
   1557     else:
-> 1558         return np.percentile(values, q, axis=axis,␣
 ↪interpolation=interpolation)
   1559
   1560

<__array_function__ internals> in percentile(*args, **kwargs)

~\anaconda3\lib\site-packages\numpy\lib\function_base.py in percentile(a, q,␣
 ↪axis, out, overwrite_input, interpolation, keepdims)
   3730     if not _quantile_is_valid(q):
   3731         raise ValueError("Percentiles must be in the range [0, 100]")
-> 3732     return _quantile_unchecked(
   3733         a, q, axis, out, overwrite_input, interpolation, keepdims)
   3734

~\anaconda3\lib\site-packages\numpy\lib\function_base.py in␣
 ↪_quantile_unchecked(a, q, axis, out, overwrite_input, interpolation, keepdims
   3849                          interpolation='linear', keepdims=False):
   3850     """Assumes that q is in [0, 1], and is an ndarray"""
```

```
-> 3851     r, k = _ureduce(a, func=_quantile_ureduce_func, q=q, axis=axis,␣
 ↪out=out,
   3852                     overwrite_input=overwrite_input,
   3853                     interpolation=interpolation)

~\anaconda3\lib\site-packages\numpy\lib\function_base.py in _ureduce(a, func,␣
 ↪**kwargs)
   3427         keepdim = (1,) * a.ndim
   3428
-> 3429     r = func(a, **kwargs)
   3430     return r, keepdim
   3431

~\anaconda3\lib\site-packages\numpy\lib\function_base.py in␣
 ↪_quantile_ureduce_func(a, q, axis, out, overwrite_input, interpolation,␣
 ↪keepdims)
   3952         weights_above.shape = weights_shape
   3953
-> 3954         ap.partition(concatenate((indices_below, indices_above)),␣
 ↪axis=axis)
   3955
   3956         # ensure axis with q-th is first

KeyboardInterrupt:
```

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[15]:
```
ap = ap.iloc[:,0:6]
ap.head(3)
len(ap)
```

[15]:
```
       id ident           type                   name  latitude_deg  \
0    6523   00A       heliport     Total Rf Heliport     40.070801
1  323361  00AA  small_airport  Aero B Ranch Airport     38.704022
2    6524  00AK  small_airport          Lowell Field     59.947733

   longitude_deg
0     -74.933601
1    -101.473911
```

```
2    -151.692524
```

[15]:
```
69063
```

[16]:
```
ap.dtypes
```

[16]:
```
id                 int64
ident             object
type              object
name              object
latitude_deg     float64
longitude_deg    float64
dtype: object
```

[36]:
```
list((ap.latitude_deg.min(),ap.latitude_deg.max(),ap.latitude_deg.mean()))
```

[36]:
```
[-90.0, 82.75, 25.986699838116575]
```

[33]:
```
ap.agg({'latitude_deg': ['min','max','mean']})
ap.agg({'latitude_deg': ['min','max','mean']}).T
```

[33]:
```
      latitude_deg
min        -90.0000
max         82.7500
mean        25.9867
```

[33]:
```
                min    max     mean
latitude_deg -90.0  82.75  25.9867
```

[59]:
```
ap.agg({'latitude_deg': ['count','min','max','mean'],'longitude_deg':␣
↪['count','min','max','mean']})
ap.agg({'latitude_deg': ['count','min','max','mean'],'longitude_deg':␣
↪['count','min','max','mean']}).T
```

[59]:
```
        latitude_deg  longitude_deg
count    69063.0000   69063.000000
min        -90.0000    -179.876999
max         82.7500     179.975700
mean        25.9867     -31.169175
```

[59]:
```
                 count         min       max       mean
latitude_deg   69063.0  -90.000000   82.7500  25.986700
longitude_deg  69063.0 -179.876999  179.9757 -31.169175
```

[60]:
```
ap.min()
```

[60]:
```
id                                    2
ident                               00A
type                        balloonport
name                "Der Dingel" Airfield
latitude_deg                        -90
longitude_deg                  -179.877
elevation_ft                      -1266
iso_region                        AD-04
scheduled_service                    no
dtype: object
```

[ ]: