OLAP queries and dashboard

Requirements

To be able to run the queries found in olap.ipynb or the dashboard you will first have to boot up and populate the postgres database. To do so follow the README.md above and make sure to have installed all requirements by running `pip install -r src/requirements.txt`.

OLAP queries

The OLAP queries are for the most part self explanatory and should be able to be seen and run by following the instructions in the notebook.

OLAP Queries Listed

1. Rollup Query: Getting the Aggregated Hectares Burnt / Cost per Year

   ```
   SELECT EXTRACT(YEAR FROM burncostdate) AS year,SUM(hectaresburnt)
   AS hectaresburnt,SUM(cost) AS cost FROM dailyburncost GROUP
   BY year ORDER BY year;
   ```

```
year |   hectaresburnt    |        cost
------+--------------------+--------------------
 1986 | 17027.881466640312 |  8799761.199425355
 1987 |  35923.44577856589 | 10032713.672953464
 1988 |   37061.5133965367 | 15754547.418136569
 1989 | 1378639.9272416614 | 327686933.20211315
 1990 |  20832.41039935276 | 10258858.663619723
 1991 |  85214.06238526855 |  36294745.85677179
 1992 |  67045.79093791095 | 11473744.399522731
 1993 |  68941.81952615915 | 7926416.2110517705
 1994 |  526351.5190959437 |  78300209.38625383
 1995 |  866634.9730330523 | 286788685.77212363
 1996 | 146752.07816767506 |   86244924.6192076
 1997 | 31904.930806261644 |  18508866.37611507
 1998 | 310891.21427160315 |  90647933.22141828
```

2. Drilldown Query: Getting Aggregated Hectares Burnt/Cost per year per Province

   ```
   SELECT EXTRACT(YEAR FROM burncostdate) AS year, fireprovinceshort
   AS province,SUM(hectaresburnt) AS hectaresburnt,SUM(cost) AS
   cost FROM dailyburncost GROUP BY GROUPING SETS ((year,province))
   ORDER BY year;
   ```

```
year | province |    hectaresburnt    |        cost
------+----------+---------------------+--------------------
 1986 | ON       |  16831.541084577115 |  8773381.188561335
```

```
1986 | SK       |  196.34038206316623 | 26380.010864011005
1987 | ON       |   14227.92792308665 |  7416257.047727833
1987 | SK       |  21695.517855478924 | 2616456.6252257414
1988 | ON       |   23462.85961072778 |  14138485.61731017
1988 | MB       |  5781.8737278200515 |  711327.2413020197
1988 | SK       |   7816.780057988726 |  904734.5595244003
1989 | NL       |  2150.3956761946715 |  884177.2590100288
1989 | MB       |   773174.7988482267 | 102631053.83335538
```

3. Slice Query: Getting all FACTs from the Year 2022

```
SELECT * FROM dailyburncost WHERE EXTRACT(YEAR FROM burncostdate)=2022
ORDER BY burncostdate
```

	stationid	burncostdate	provinceid	burnincidentid	fireprovinceshort	averagetemperature	maxrelativehumidity	maxwindspeedgust	hectaresburnt	cost
0	3608	2022-02-13	8	257170	AB	-2.0	100.0	31.0	0.050305	71.152275
1	3375	2022-03-23	8	256865	AB	11.5	NaN	NaN	0.094233	133.283341
2	635	2022-04-23	9	252012	BC	9.9	96.0	NaN	0.049156	198.651495
3	3422	2022-04-24	8	251337	AB	6.2	94.0	37.0	0.886300	1253.589332
4	5690	2022-04-30	5	254631	ON	7.2	79.0	NaN	0.639585	4649.222251
...	...	...	...	...	...	...	...	...	...	...
5616	1128	2022-11-18	9	249370	BC	-3.8	79.0	37.0	0.307776	1243.809830
5617	1663	2022-11-18	9	252643	BC	-15.2	89.0	NaN	74.673126	301774.922711
5618	1663	2022-11-19	9	252644	BC	-10.9	91.0	NaN	74.673126	301774.922711
5619	1663	2022-11-20	9	252645	BC	-6.5	95.0	NaN	74.673126	301774.922711
5620	1663	2022-11-21	9	252646	BC	-1.1	95.0	33.0	74.673126	301774.922711

5621 rows × 10 columns

Figure 1: text

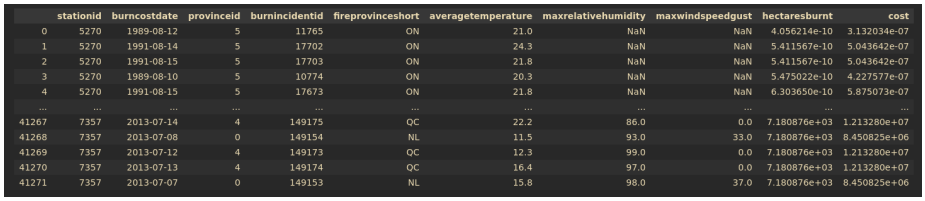
4. Dice Query: Getting the per Province Per Year Aggregated Hectares Burnt and Costs for the Provinces of Quebec and British Columbia

```
SELECT EXTRACT(YEAR FROM burncostdate) AS year,fireprovinceshort
AS province,SUM(hectaresburnt) AS hectaresburnt,SUM(cost)
AS cost FROM dailyburncost WHERE fireprovinceshort='QC' OR
fireprovinceshort='BC' GROUP BY GROUPING SETS ((year,fireprovinceshort))
ORDER BY year;
```

```
 year | province |    hectaresburnt    |        cost
------+----------+---------------------+--------------------
 1989 | BC       |     801.22030819001 |  311927.9386802062
 1989 | QC       |   417050.8967506842 | 131815762.99566288
 1990 | BC       |  1200.2578239916022 |  526053.9147644294
 1990 | QC       |   1531.475724106538 |  532390.7919900914
 1991 | QC       |  35161.019537605025 | 13062653.150757343
 1991 | BC       |   939.3577084399344 |  452381.8992486935
 1992 | QC       |  113.12012299568964 | 43169.678666505606
 1992 | BC       |    2129.72786970193 | 1070466.2137473891
 1993 | QC       |    368.954055574313 | 145879.44649781616
 1993 | BC       | 0.08747862706606391 |  49.52755939515313
 1994 | QC       |   3840.322969722674 | 1602336.2937208025
 1994 | BC       |   616.5227659669126 |  396460.8155084678
 1995 | BC       |  1483.4053860478775 | 1060775.9276190195
```

5. Dice Query 2: Getting all FACTs where there is greater than 200 hectares burnt or average temp is greater than 20

```
SELECT * FROM dailyburncost WHERE hectaresburnt>200.0 OR
averagetemperature > 20 ORDER BY hectaresburnt ASC;
```



	stationid	burncostdate	provinceid	burnincidentid	fireprovinceshort	averagetemperature	maxrelativehumidity	maxwindspeedgust	hectaresburnt	cost
0	5270	1989-08-12	5	11765	ON	21.0	NaN	NaN	4.056214e-10	3.132034e-07
1	5270	1991-08-14	5	17702	ON	24.3	NaN	NaN	5.411567e-10	5.043642e-07
2	5270	1991-08-15	5	17703	ON	21.8	NaN	NaN	5.411567e-10	5.043642e-07
3	5270	1989-08-10	5	10774	ON	20.3	NaN	NaN	5.475022e-10	4.227577e-07
4	5270	1991-08-15	5	17673	ON	21.8	NaN	NaN	6.303650e-10	5.875073e-07
...	...	...	...	...	...	...	...	...	...	...
41267	7357	2013-07-14	4	149175	QC	22.2	86.0	0.0	7.180876e+03	1.213280e+07
41268	7357	2013-07-08	0	149154	NL	11.5	93.0	33.0	7.180876e+03	8.450825e+06
41269	7357	2013-07-12	4	149173	QC	12.3	99.0	0.0	7.180876e+03	1.213280e+07
41270	7357	2013-07-13	4	149174	QC	16.4	97.0	0.0	7.180876e+03	1.213280e+07
41271	7357	2013-07-07	0	149153	NL	15.8	98.0	37.0	7.180876e+03	8.450825e+06

Figure 2: text

6. Comb Query 1: Getting average cost and hectares burned from FACT Table for fires that have burned more than 10 hectares grouped by province (Rollup + Slice)

```
SELECT fireprovinceshort, AVG(cost) as cost,AVG(hectaresburnt)
as avg_burnt, COUNT(burnincidentid) as fires FROM dailyburncost
WHERE hectaresburnt > 10 GROUP BY fireprovinceshort
```

fireprovinceshort	cost	avg_burnt	fires
NL	354559.55060377024	306.10658650885557	571
NS	64389.88283441146	66.64730327994103	10
QC	163011.24117769834	200.98308753362286	6750
ON	253802.0590604995	93.87534816590585	9739
MB	61366.15314534975	160.53798155974187	14950
SK	51334.11666419762	154.95906082830376	18304
AB	70249.51028253877	102.4188332134256	6006
BC	302068.7708467641	111.05944081250921	9818

7. Comb Query 2: Getting average Temp, Hectares Burnt, and Total cost from fires burning more than 10 hectares by year. (Rollup + Slice)

```
SELECT EXTRACT(YEAR FROM burncostdate) as year, AVG (averagetemperature)
as avg_temp, AVG(hectaresburnt) as avg_burnt, SUM(cost) as
total_cost FROM dailyburncost WHERE hectaresburnt > 10 GROUP
BY year ORDER BY year;
```

year	avg_temp	avg_burnt	total_cost
1986	15.065053763440854	84.47487399176363	8189994.582136311
1987	16.13895747599444	44.88821023212052	8684482.33694823
1988	16.056714628297343	34.95451087265067	11583502.368882671
1989	17.912069756769522	314.3479198833328	323865507.15756345
1990	15.83274111675128	45.42988452887317	8768303.96628786

```
1991 | 17.467824497257673 |  76.43773029026994 |  35570380.01278175
1992 | 14.533806146572124 |  155.7814339868413 | 10889832.398385208
1993 |  12.69942196531793 | 194.76202004253352 |  7470136.468116514
1994 | 15.614835747086047 | 184.43397109467898 |   77155473.3159063
1995 | 16.480870396939448 | 205.92798774311274 |  283805284.2387678
1996 |  16.08126326963902 |  74.63139936154586 |  82653252.56108153
1997 | 14.141843971631243 | 52.883526889490824 |  17446497.18393898
```

8. Comb Query 3: Compare provinces in term of average temperature and total cost and hecctares burnt looking only at the month of August.

```
SELECT fireprovinceshort, AVG(averagetemperature) as avgtemp
,SUM(hectaresburnt) as totalburnt,SUM(cost) as totalcost
FROM dailyburncost \  WHERE EXTRACT(MONTH FROM burncostdate)=8
GROUP BY fireprovinceshort ORDER BY avgtemp DESC
```

```
 fireprovinceshort    avgtemp        totalburnt      totalcost
-------------------+-------------+---------------+-----------------------
                NS | 19.592157   |     53.080946 | 4.502069e+04
                NB | 18.179186   |      2.906608 | 2.152975e+03
                BC | 17.905035   | 611948.569143 | 1.612825e+09
                MB | 16.695415   | 816247.783700 | 2.933066e+08
                ON | 16.446630   | 227398.673711 | 5.800051e+08
                SK | 16.287687   | 550868.924205 | 2.039902e+08
                QC | 16.017258   | 255365.486559 | 1.109103e+08
                AB | 15.742508   |  90858.083114 | 6.322530e+07
                NL | 15.433447   |  12731.810064 | 1.434929e+07
```

9. Comb Query 4: Getting fires in June, July, or August in Ontario by year (Slice + rollup)

```
SELECT EXTRACT(YEAR FROM burncostdate) as year, AVG(averagetemperature)
as avg_temp,SUM(hectaresburnt) as total_hectare_burnt,SUM(cost)
as cost FROM (SELECT * FROM dailyburncost WHERE EXTRACT(MONTH
FROM burncostdate)=6 OR EXTRACT(MONTH FROM burncostdate)=7
OR EXTRACT(MONTH FROM burncostdate)=8 ) WHERE fireprovinceshort='ON'
GROUP BY year ORDER BY year ASC;
```

```
 year |      avg_temp      | total_hectare_burnt |        cost
------+--------------------+---------------------+--------------------
 1986 | 14.469951534733449 |  13564.149788102195 |  7070265.045356327
 1987 | 16.545919778699897 |   13227.67173394098 |  6894877.051119337
 1988 | 18.000817307692262 |  17475.794743917017 | 10530739.932696287
 1989 | 18.220064550833794 |  106844.98320244312 |  82501103.98634677
 1990 | 16.400829875518685 |    9200.85485524539 |  7994429.532258958
 1991 |  17.58764044943821 |   19651.12095414346 | 18315067.404853284
 1992 | 12.824260355029585 |  1637.9818507074642 | 1447734.6669142374
 1993 | 15.806329113924049 |   654.2675861598867 |  567685.0282180471
 1994 | 17.094249201277954 |  1252.5188646774718 |  1081697.797337804
```

```
 1995 | 18.146010844306755 |   137297.0886663584 | 121572654.79643449
 1996 | 15.859498480243149 |  36503.620479258956 | 35218385.763882056
 1997 | 15.334394904458597 |   910.0298040610726 |   983686.607303572
 1998 |  16.07062975027142 |  11813.211204389007 | 13448224.653155047
 1999 | 14.621813725490206 |  11751.537601831027 | 13791768.128413608
 2000 |  16.29876543209876 |  315.16228867761544 |  378041.6930960958
 2001 |  17.22814465408806 |   456.5877039496165 |  559509.3353134096
```

10. Iceberg Query: Getting the 10 weeks of the year with the most fires, as well as their average temperature and average costs

```
SELECT EXTRACT(WEEK FROM burncostdate) as week, AVG(averagetemperature)
avg_temp, COUNT(burnincidentid) AS fires, AVG(cost) as
avg_cost FROM dailyburncost GROUP BY week ORDER BY fires
DESC LIMIT 10;
```

```
 week |      avg_temp      | fires |      avg_cost
------+--------------------+-------+--------------------
   29 |  17.75540386038046 | 14869 |  67915.83662463122
   28 | 17.535134949522444 | 14561 |  84718.62231880634
   30 | 17.739053377814695 | 14388 | 57222.219177688006
   31 |   17.2993579225225 | 14017 |  55380.37973021749
   27 | 16.919980830199563 | 13563 |  76063.62693145222
   32 | 17.460191791330914 | 13035 | 62704.451023903835
   33 | 16.882128216310335 | 11465 |  56814.51364408961
   26 |  16.58846457037751 | 11417 | 45434.346044547034
   25 | 15.653364900262055 |  9926 | 34526.378918785165
   34 | 15.785639508420681 |  8788 | 51673.873800171146
(10 rows)
```

11. Windowing Query: A ranking of months by most hectares burnt for each year

```
SELECT EXTRACT(YEAR FROM burncostdate) as year, EXTRACT(MONTH
FROM burncostdate) as month,AVG(averagetemperature) as
avg_temp,SUM(hectaresburnt) as total_hectare_burnt,SUM(cost)
as cost, RANK() OVER (PARTITION BY EXTRACT(YEAR FROM burncostdate)
ORDER BY SUM(hectaresburnt) DESC) FROM dailyburncost GROUP
BY GROUPING SETS ((year,month)) ORDER BY year,rank;
```

```
year | month |       avg_temp       | total_hectare_burnt  |        cost        | rank
------+-------+----------------------+----------------------+--------------------+------
 1986 |     6 |   12.967687074829936 |      9445.2399858812 |  4890738.601219509 |    1
 1986 |     7 |   16.335377358490565 |   3971.0091406412243 | 2065723.4363922821 |    2
 1986 |     5 |   16.169620253164553 |   3185.5047963637553 | 1653527.8395789892 |    3
 1986 |     8 |   15.026203208556154 |   304.04427838601526 | 134782.23914061638 |    4
 1986 |     9 |   7.6927536231884055 |    87.11618064104631 | 45409.000674086536 |    5
 1986 |     4 |                  3.9 |     22.3485580217033 | 3002.7200579544706 |    6
 1986 |    10 |   2.8166666666666664 |   12.618526705120624 |  6577.362361875905 |    7
```

```
 1987 |     7 |    16.96859838274926 |   16441.274756242743 |   4658313.32717081 |    1
 1987 |     6 |   16.010196078431413 |   12191.896422823502 | 3176618.8636521394 |    2
 1987 |     5 |   10.720469798657735 |   3739.1830863782748 |  828913.8189015682 |    3
 1987 |     8 |   14.623546511627946 |    3487.571637192051 | 1338429.3534290253 |    4
 1987 |     4 |     8.95416666666667 |   34.191076191792334 |  15150.77679246737 |    5
```

12. Usage of Window Clause: A calculation of the percentage of hectares burnt in a 3 month moving average

```
SELECT burnincidentid,EXTRACT(YEAR FROM burncostdate) as
year,EXTRACT(MONTH FROM burncostdate) as month,hectaresburnt,
hectaresburnt / SUM(hectaresburnt) OVER W AS fires_mov_avg
FROM dailyburncost GROUP BY GROUPING SETS ((burnincidentid,year,month,hectaresburnt))
WINDOW W AS (PARTITION BY EXTRACT(YEAR FROM burncostdate)
ORDER BY EXTRACT(MONTH FROM burncostdate) RANGE BETWEEN '1'
PRECEDING AND '1' FOLLOWING)
```

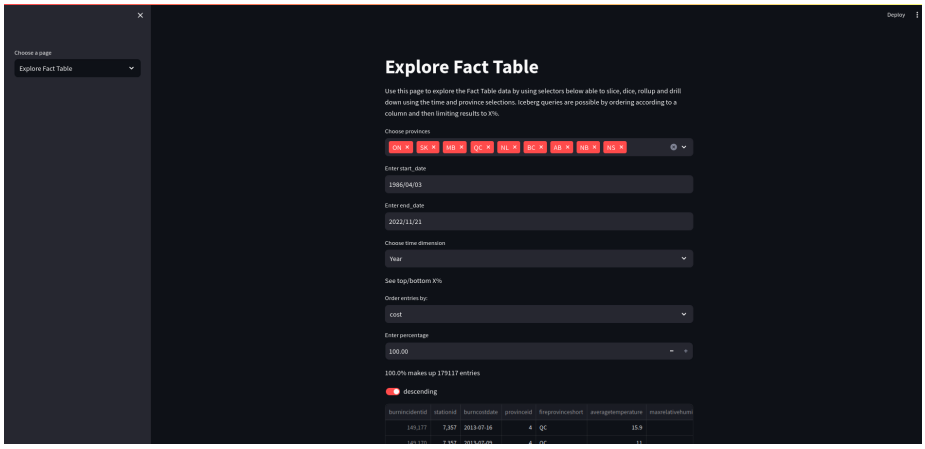
```
 burnincidentid | year | month |     hectaresburnt      |     fires_mov_avg
----------------+------+-------+------------------------+-----------------------
            188 | 1986 |     4 |       4.46971160434066 |  0.0013933653164756165
            189 | 1986 |     4 |       4.46971160434066 |  0.0013933653164756165
            190 | 1986 |     4 |       4.46971160434066 |  0.0013933653164756165
            186 | 1986 |     4 |       4.46971160434066 |  0.0013933653164756165
            187 | 1986 |     4 |       4.46971160434066 |  0.0013933653164756165
            274 | 1986 |     5 |     1.0265254301647333 |  8.112841678785102e-05
            273 | 1986 |     5 |     1.0265254301647333 |  8.112841678785102e-05
            109 | 1986 |     5 |     11.745024203524355 |  0.0009282334278012099
            110 | 1986 |     5 |     11.745024203524355 |  0.0009282334278012099
            111 | 1986 |     5 |     11.745024203524355 |  0.0009282334278012099
            112 | 1986 |     5 |     11.745024203524355 |  0.0009282334278012099
            548 | 1986 |     5 |     4.7396281176610975 |   0.000374582561765977
            549 | 1986 |     5 |     4.7396281176610975 |   0.000374582561765977
            272 | 1986 |     5 |     1.0265254301647333 |  8.112841678785102e-05
```

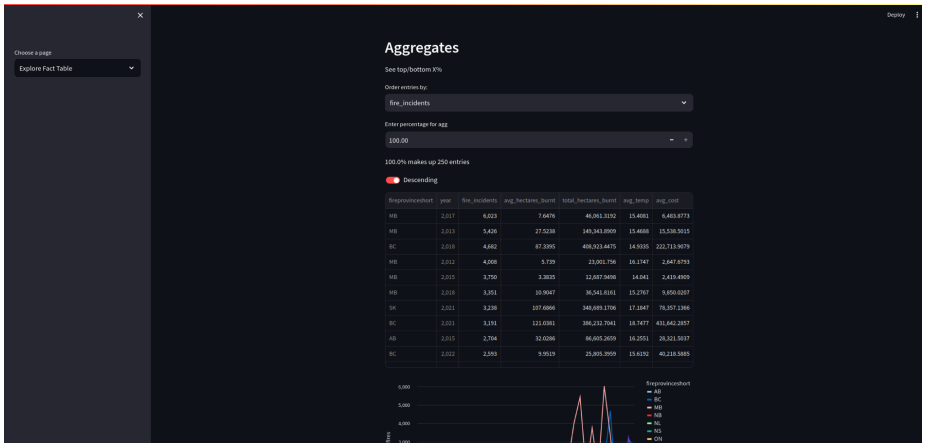
Dashboard

For the dashboard you will have to create a file at '~/.streamlit/secrets.toml' with contents in this format (Consult Streamlit documentation for location on Windows): `[connections.postgresql] dialect = "postgresql" host = "localhost" port = "5432" database = "firedb" username = "username" password = "password"`

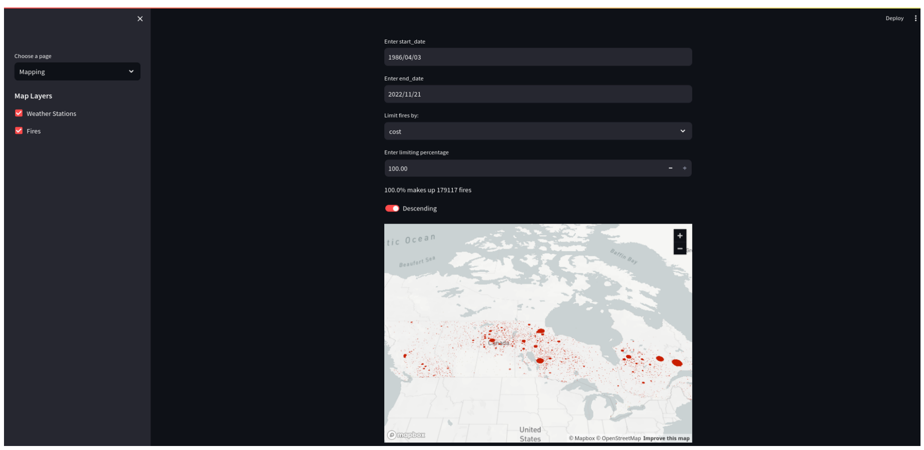
The dashboard is produced using the library Streamlit which produces an interactable dashboard in the browser based on the scripting in dashboard.py. To start it up, ensure all requirements are installed and run `streamlit run dashboard.py`. The start page is mostly empty so navigate with the sidebar to one of the two other pages to begin exploring the dashboard. One thing to note is that "Time-dimension" means on which timescale the aggregates are done,

where “Year-Month” treats all months across the years seperately while “Month” groups all fires in e.g. March together no matter the year. Some examples of the dashboard are produced below for your convenience:






Choose a page
Explore Fact Table
Deploy
fires
time
avg_hectares_burnt
fireprovinceshort
AB
BC
MB
NB
NL
NS
ON
QC
SK
total_hectares_burnt


Choose a page
Mapping
Map Layers
Weather Stations
Fires
Deploy
Enter start_date
1986/04/03
Enter end_date
2022/11/21
Limit fires by:
cost
Enter limiting percentage
100.00
100.0% makes up 179117 fires
Descending
United States
© Mapbox © OpenStreetMap Improve this map