PROJECT PART 4

Queries:

1. WRITE A QUERY TO DISPLAY different statecodes and sourcenames whose productId is 38344AK060.

```
select distinct ("StateCode","SourceName")

from businessrules

where "ProductId" = '38344AK060';
```

2. WRITE A QUERY TO DISPLAY names that are repeated

```
Select "BenefitName" ,

count(*) from "benefitcostsharing"

group by "BenefitName"

having count(*) >1;
```

3.WRITE A QUERY TO DISPLAY total and average of version numbers

```
select sum("VersionNum"),avg("VersionNum")

from rate where "BussinessYear" ='2014';
```

4. WRITE A QUERY TO DISPLAY maximum, minimum, total, and number of version numbers whose state code is 'AK' and having IssuerId greater that 30000

```
select max("VersionNum"),min("VersionNum"),

avg("VersionNum"),sum("VersionNum"),count("VersionNum")

from servicearea where "StateCode" ='AK' and "IssuerId" > '30000';
```

5. WRITE A QUERY TO DISPLAY sum of IssuerId,IssuerId2 is greater that 800000 and Version number greater than 7

```
select "IssuerId","IssuerId2"

from servicearea

where ("IssuerId" + "IssuerId2")>'80000'
```

and "VersionNum" > '7';

6. WRITE A QUERY TO DISPLAY subtraction of IssuerId,IssuerId2 is Null and VersionNum should be null

```
Select "IssuerId" , "IssuerId2" from rate

where ("IssuerId" - "IssuerId2")='0'

and "VersionNum" = '9';
```

7.WRITE A QUERY TO DISPLAY number of Statecode that are repeated more than thrice having sourcename 'HIOS'

```
select "StateCode" , count(*)

from servicearea

where "SourceName" ='HIOS'

group by "StateCode"

having count(*) > '3';
```

8.WRITE A QUERY TO DISPLAY number of VersionNum that are repeated more than 5 having Statecode 'AK'

```
select "VersionNum", count(*)

from rate where "StateCode" = 'AK'

group by "VersionNum" having count(*) > '5';
```

9.WRITE A QUERY TO DISPLAY Whose Benefitname begin with 'R' or 'M'

```
select "BenefitName"

from benefitcostsharing

where "BenefitName" like '%R' or "BenefitName" like '%M';
```

QUERY 1

Query | Query History

```
1  select distinct ("StateCode","SourceName")
2  from businessrules
3  where "ProductId" = '38344AK060';
```

Data Output | Messages | Notifications

| | row record |
|---|---|
| 1 | (AK,HIOS) |
| 2 | (AK,OPM) |

```
def runQuery(conn):
    select_Query =  "select distinct (\"StateCode\",\"SourceName\") from businessrules where \"ProductId\" = '38344AK060';"
    servicearea_df = pd.DataFrame(columns = ['StateCode','SourceName'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'StateCode': row[0],'SourceName': row[1]}
            servicearea_df = servicearea_df.append(output_df, ignore_index=True)

        print(servicearea_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
  StateCode SourceName
0        AK       HIOS
1        AK        OPM

= RESTART: C:/Users/student/AppData/Local/Programs/Python/Python311/connect.py
```

QUERY 2

Query | Query History

```
1  Select "BenefitName" ,
2  count(*) from "benefitcostsharing"
3  group by "BenefitName"
4  having count(*) >1;
```

Data Output | Messages | Notifications

| | BenefitName character varying | count bigint |
|---|---|---|
| 1 | Routine Exams a... | 8 |
| 2 | ΞClinical - TrialsΞ | 55 |
| 3 | ΞDiabetes - Care ... | 55 |
| 4 | ΞMental Health - ... | 55 |
| 5 | ΞPrescription Dru... | 55 |
| 6 | 24/7 Doctor Help... | 24 |
| 7 | ABA Autism Spe... | 4 |

```
def runQuery(conn):
    select_Query =  "Select \"BenefitName\" , count(*) from benefitcostsharing group by \"BenefitName\" having count(*) >1;"
    benefitcostsharing_df = pd.DataFrame(columns = ['BenefitName'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'BenefitName': row[0],'count': row[1]}
            benefitcostsharing_df = benefitcostsharing_df.append(output_df, ignore_index=True)

        print(benefitcostsharing_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
FutureWarning: The frame.append method is deprecated and will
                                  BenefitName    count
0           Routine Exams and X-Rays - Child      8.0
1                      Clinical - Trials     55.0
2               Diabetes - Care Management     55.0
3                    Mental Health - Other     55.0
4               Prescription Drugs - Other     55.0
..                                     ...      ...
851                          Wilm's Tumor    988.0
852            X-rays and Diagnostic Imaging  65704.0
853           X-rays of Entire Mouth - Adult     20.0
854           X-rays of Entire Mouth - Child     32.0
855  Xrays and Diagnostic Imaging (Hospital)     48.0

[856 rows x 2 columns]
```

QUERY 3

Query  Query History

```
1  select sum("VersionNum"),avg("VersionNum")
2  from rate where "BussinessYear" ='2014';
```

Data Output  Messages  Notifications

| | sum bigint | avg numeric |
|---|---|---|
| 1 | 26206262 | 6.902946168 |

```
def runQuery(conn):
    select_Query =  "select sum(\"VersionNum\"),avg(\"VersionNum\") from rate where \"BussinessYear\" ='2014';"
    rate_df = pd.DataFrame(columns = ['VersionNum'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'Sum': row[0],'Avg': row[1]}
            rate_df = rate_df.append(output_df, ignore_index=True)

        print(rate_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
   VersionNum          Sum                Avg
0         NaN  26206262.0  6.9029461688320583
```

QUERY 4

Query | Query History

```
1  select max("VersionNum"),min("VersionNum"),
2  avg("VersionNum"),sum("VersionNum"),count("VersionNum")
3  from servicearea where "StateCode" ='AK' and "IssuerId" > '30000';
```

Data Output | Messages | Notifications

| | max integer | min integer | avg numeric | sum bigint | count bigint |
|---|---|---|---|---|---|
| 1 | 15 | 1 | 4.4571428571428571 | 156 | 35 |

```
def runQuery(conn):
    select_Query =  "select max(\"VersionNum\"),min(\"VersionNum\"), avg(\"VersionNum\"),sum(\"VersionNum\"),count(\"VersionNum\")
                    from servicearea where \"StateCode\" ='AK' and \"IssuerId\" > '30000';;"
    servicearea_df = pd.DataFrame(columns = ['VersionNum'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'max': row[0],'min': row[1],'avg': row[2],'sum': row[3],'count': row[4]}
            servicearea_df = servicearea_df.append(output_df, ignore_index=True)

        print(servicearea_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
   VersionNum   max  min                  avg    sum  count
0         NaN  15.0  1.0  4.4571428571428571  156.0   35.0
```

QUERY 5

Query  Query History

```
1  select "IssuerId","IssuerId2"
2  from servicearea
3  where ("IssuerId" + "IssuerId2")>'80000'
4  and "VersionNum" > '7';
```

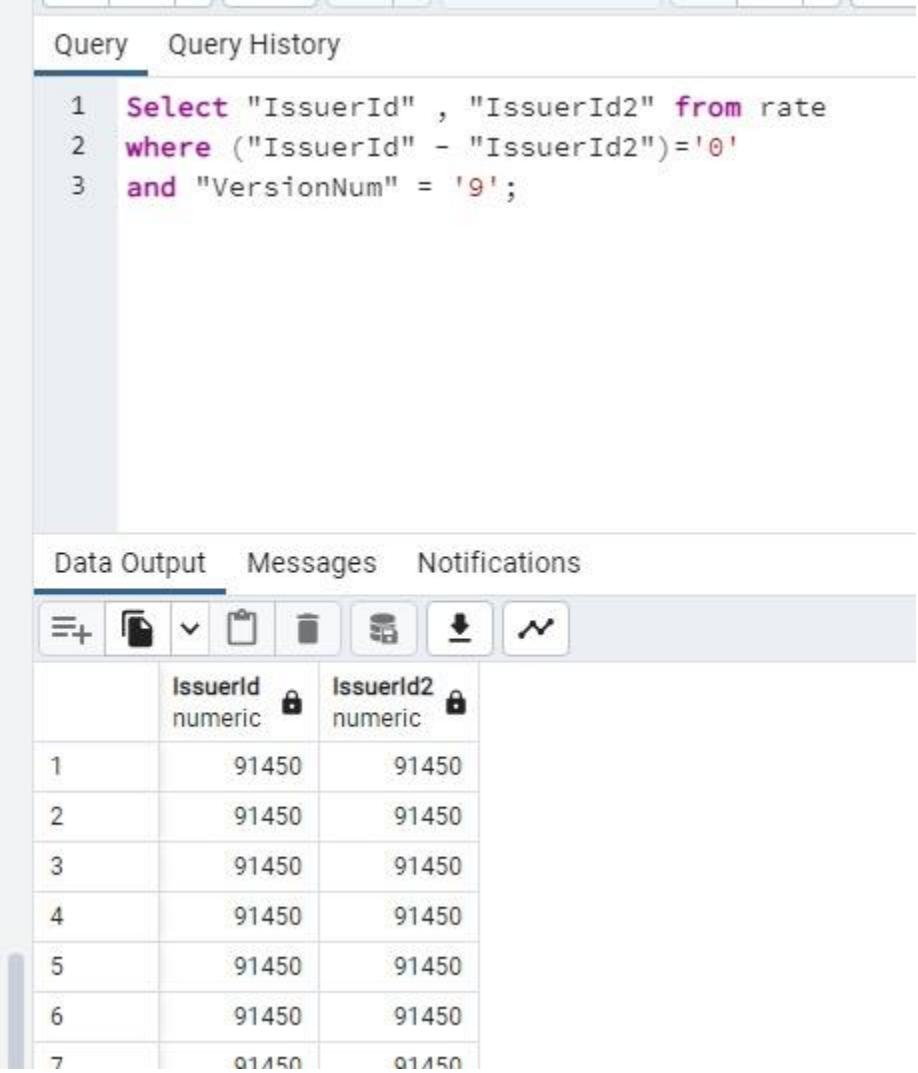
Data Output  Messages  Notifications

| | IssuerId integer | IssuerId2 integer |
|---|---|---|
| 1 | 44580 | 44580 |
| 2 | 44580 | 44580 |
| 3 | 44580 | 44580 |
| 4 | 44580 | 44580 |
| 5 | 59809 | 59809 |
| 6 | 51485 | 51485 |
| 7 | 70239 | 70239 |

```
def runQuery(conn):
    select_Query =  "select \"IssuerId\",\"IssuerId2\" from servicearea where (\"IssuerId\" + \"IssuerId2\")>'80000' and \"VersionNum\" > '7';"
    servicearea_df = pd.DataFrame(columns = ['IssuerId','IssuerId2'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'IssuerId': row[0],'IssuerId2': row[1]}
            servicearea_df = servicearea_df.append(output_df, ignore_index=True)

        print(servicearea_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
      IssuerId IssuerId2
0        44580     44580
1        44580     44580
2        44580     44580
3        44580     44580
4        59809     59809
...        ...       ...
8240     99663     99663
8241     99663     99663
8242     99663     99663
8243     99663     99663
8244     99663     99663

[8245 rows x 2 columns]
```

QUERY 6




QUERY 7

Query | Query History

```
1  select "VersionNum", count(*)
2  from rate where "StateCode" = 'AK'
3  group by "VersionNum" having count(*) > '5';
```

Data Output | Messages | Notifications

| | VersionNum integer | count bigint |
|---|---|---|
| 1 | 1 | 2760 |
| 2 | 2 | 13308 |
| 3 | 3 | 1680 |
| 4 | 4 | 22060 |
| 5 | 6 | 8424 |
| 6 | 7 | 3060 |
| 7 | 8 | 138 |

Total rows: 9 of 9 Query complete 00:00:02.909

```
def runQuery(conn):
    select_Query =  "select \"VersionNum\", count(*) from rate where \"StateCode\" = 'AK' group by \"VersionNum\" having count(*) > '5';"
    rate_df = pd.DataFrame(columns = ['VersionNum'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'VersionNum': row[0],'count': row[1]}
            rate_df = rate_df.append(output_df, ignore_index=True)

        print(rate_df)

def main():
    conn = initialize()
    runQuery(conn)
```

| | VersionNum | count |
|---|---|---|
| 0 | 1 | 2760.0 |
| 1 | 2 | 13308.0 |
| 2 | 3 | 1680.0 |
| 3 | 4 | 22060.0 |
| 4 | 6 | 8424.0 |
| 5 | 7 | 3060.0 |
| 6 | 8 | 138.0 |
| 7 | 10 | 9246.0 |
| 8 | 15 | 23736.0 |

QUERY 8

Query | Query History

```
1  select "VersionNum", count(*)
2  from rate where "StateCode" = 'AK'
3  group by "VersionNum" having count(*) > '5';
```

Data Output | Messages | Notifications

| | VersionNum integer | count bigint |
|---|---|---|
| 1 | 1 | 2760 |
| 2 | 2 | 13308 |
| 3 | 3 | 1680 |
| 4 | 4 | 22060 |
| 5 | 6 | 8424 |
| 6 | 7 | 3060 |
| 7 | 8 | 138 |

Total rows: 9 of 9 | Query complete 00:00:02.909

```
def runQuery(conn):
    select_Query =  "select \"VersionNum\", count(*) from rate where \"StateCode\" = 'AK' group by \"VersionNum\" having count(*) > '5';"
    rate_df = pd.DataFrame(columns = ['VersionNum'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'VersionNum': row[0],'count': row[1]}
            rate_df = rate_df.append(output_df, ignore_index=True)

        print(rate_df)

def main():
    conn = initialize()
    runQuery(conn)
```

| | VersionNum | count |
|---|---|---|
| 0 | 1 | 2760.0 |
| 1 | 2 | 13308.0 |
| 2 | 3 | 1680.0 |
| 3 | 4 | 22060.0 |
| 4 | 6 | 8424.0 |
| 5 | 7 | 3060.0 |
| 6 | 8 | 138.0 |
| 7 | 10 | 9246.0 |
| 8 | 15 | 23736.0 |

QUERY 9

Query | Query History

```
1  select "BenefitName"
2  from benefitcostsharing
3  where "BenefitName" like '%R' or "BenefitName" like '%M';
```

Data Output | Messages | Graph Visualiser | Notifications

| | BenefitName character varying |
|---|---|
| 1 | Psychiatric Asse... |
| 2 | Psychiatric Asse... |
| 3 | Psychiatric Asse... |
| 4 | Psychiatric Asse... |
| 5 | Psychiatric Asse... |
| 6 | Psychiatric Asse... |

```
def runQuery(conn):
    select_Query =  "select \"BenefitName\"  from benefitcostsharing where \"BenefitName\" like '%R' or \"BenefitName\" like '%M';"
    benefitcostsharing_df = pd.DataFrame(columns = ['BenefitName'])
    with conn.cursor() as cursor:
        cursor.execute(select_Query)
        records = cursor.fetchall()

        for row in records:
            output_df = {'BenefitName': row[0]}
            benefitcostsharing_df = benefitcostsharing_df.append(output_df, ignore_index=True)

        print(benefitcostsharing_df)

def main():
    conn = initialize()
    runQuery(conn)
```

```
                                    BenefitName
0   Psychiatric Assessment and Stabilization in ER
1   Psychiatric Assessment and Stabilization in ER
2   Psychiatric Assessment and Stabilization in ER
3   Psychiatric Assessment and Stabilization in ER
4   Psychiatric Assessment and Stabilization in ER
5   Psychiatric Assessment and Stabilization in ER
6   Psychiatric Assessment and Stabilization in ER
7   Psychiatric Assessment and Stabilization in ER
8   Psychiatric Assessment and Stabilization in ER
9   Psychiatric Assessment and Stabilization in ER
10  Psychiatric Assessment and Stabilization in ER
11  Psychiatric Assessment and Stabilization in ER
12  Psychiatric Assessment and Stabilization in ER
13  Psychiatric Assessment and Stabilization in ER
14  Psychiatric Assessment and Stabilization in ER
15  Psychiatric Assessment and Stabilization in ER
16  Psychiatric Assessment and Stabilization in ER
```

Query comparison:

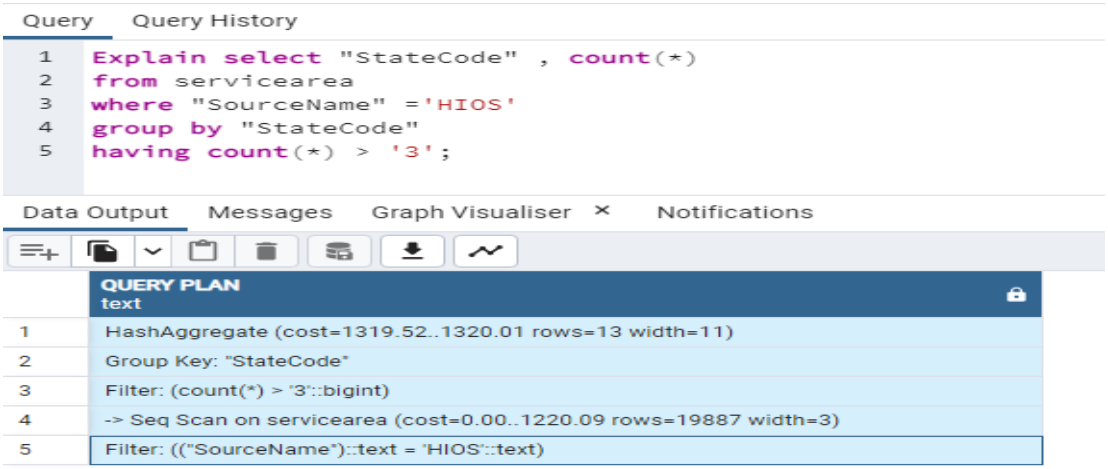
1. select "StateCode" , count(*)

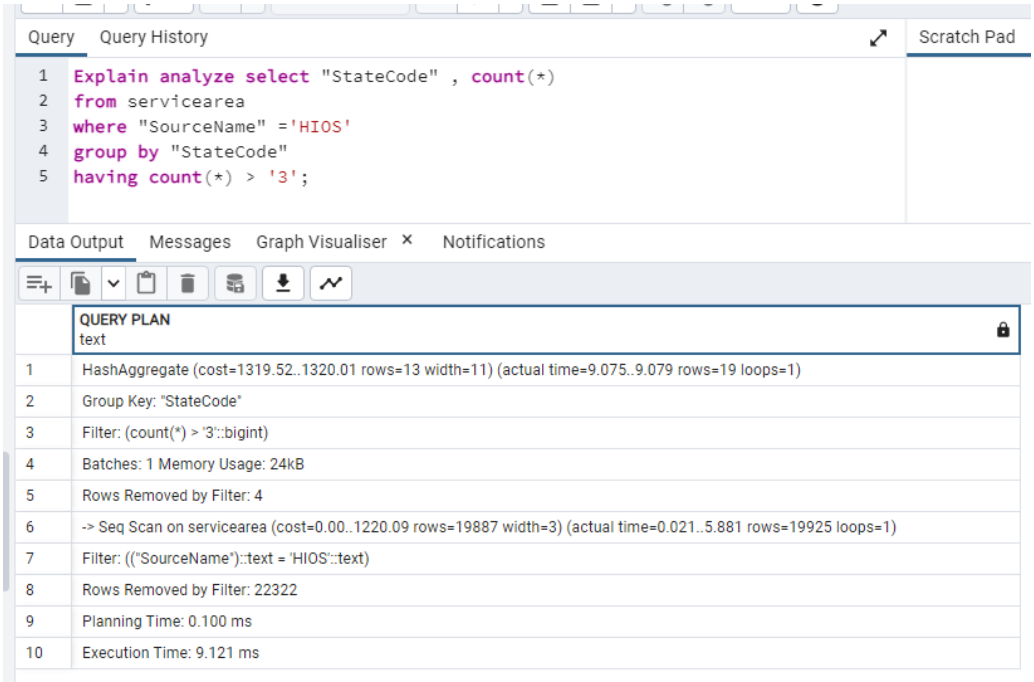
from servicearea

where "SourceName" ='HIOS'

group by "StateCode"

having count(*) > '3';



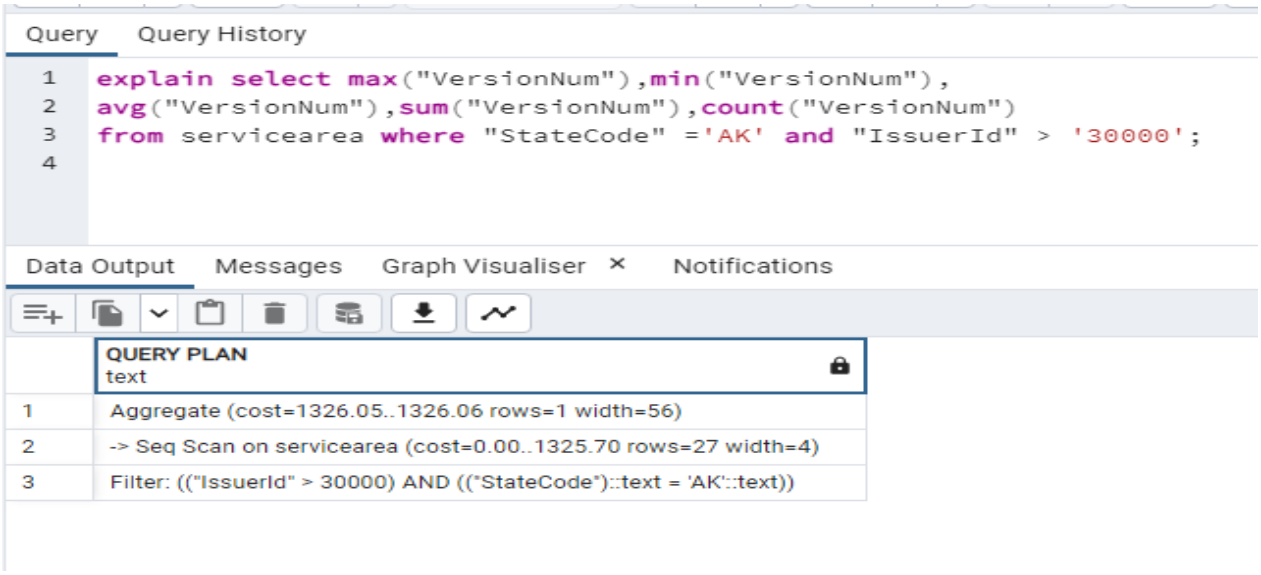



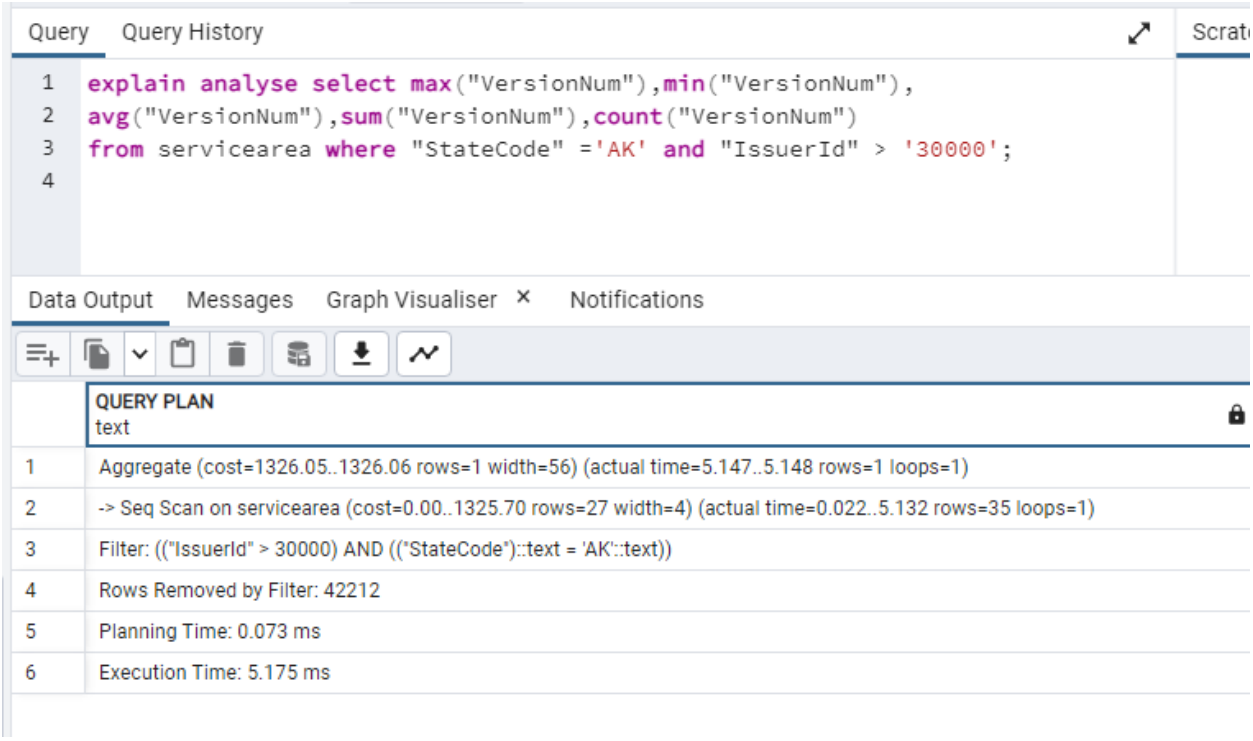

2.

select max("VersionNum"),min("VersionNum"),

avg("VersionNum"),sum("VersionNum"),count("VersionNum")

from servicearea where "StateCode" ='AK' and "IssuerId" > '30000';




3. select "BenefitName"

from benefitcostsharing

where "BenefitName" like '%R' or "BenefitName" like '%M';

Query  Query History

```
1  explain select "BenefitName"
2  from benefitcostsharing
3  where "BenefitName" like '%R' or "BenefitName" like '%M';
```

Data Output  Messages  Graph Visualiser  Notifications

| | QUERY PLAN text |
|---|---|
| 1 | Gather (cost=1000.00..207145.61 rows=19 width=28) |
| 2 | Workers Planned: 2 |
| 3 | -> Parallel Seq Scan on benefitcostsharing (cost=0.00..206143.71 rows=8 width=28) |
| 4 | Filter: ((("BenefitName")::text ~~ '%R'::text) OR (("BenefitName")::text ~~ '%M'::text)) |

Query  Query History  Scratch Pad

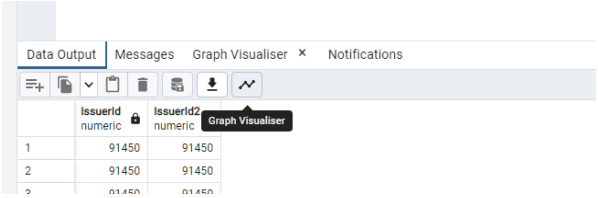
```
1  explain analyse select "BenefitName"
2  from benefitcostsharing
3  where "BenefitName" like '%R' or "BenefitName" like '%M';
```

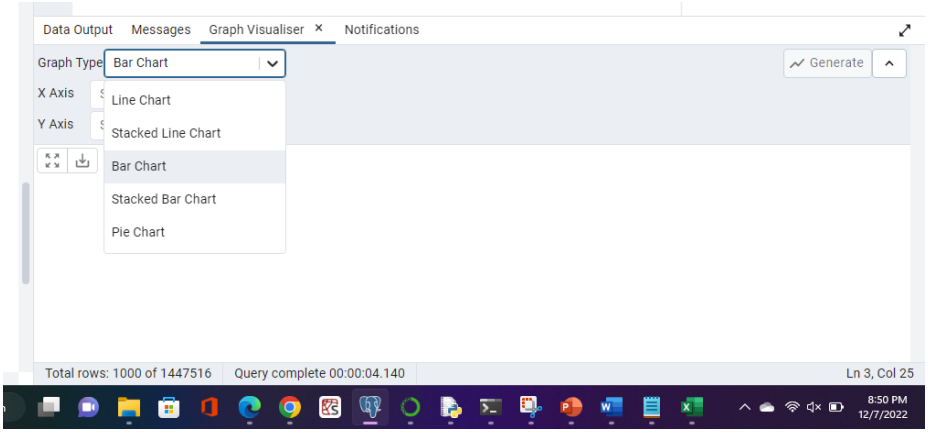
Data Output  Messages  Graph Visualiser  Notifications

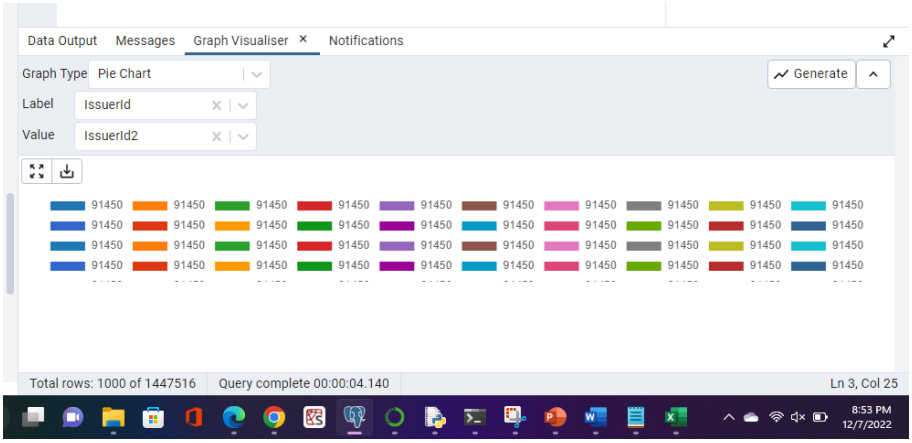
| | QUERY PLAN text |
|---|---|
| 1 | Gather (cost=1000.00..207145.61 rows=19 width=28) (actual time=2325.749..2344.300 rows=17 loops=1) |
| 2 | Workers Planned: 2 |
| 3 | Workers Launched: 2 |
| 4 | -> Parallel Seq Scan on benefitcostsharing (cost=0.00..206143.71 rows=8 width=28) (actual time=1745.892..2280.341 rows=6 loops=3) |
| 5 | Filter: ((("BenefitName")::text ~~ '%R'::text) OR (("BenefitName")::text ~~ '%M'::text)) |
| 6 | Rows Removed by Filter: 1682797 |
| 7 | Planning Time: 0.204 ms |
| 8 | Execution Time: 2344.323 ms |

Data Visualization

We will have the choice to graph visualize our data in several ways in the PG Admin; we may get the bar graph, line graph, pie graph, etc., of the data we wish to display.

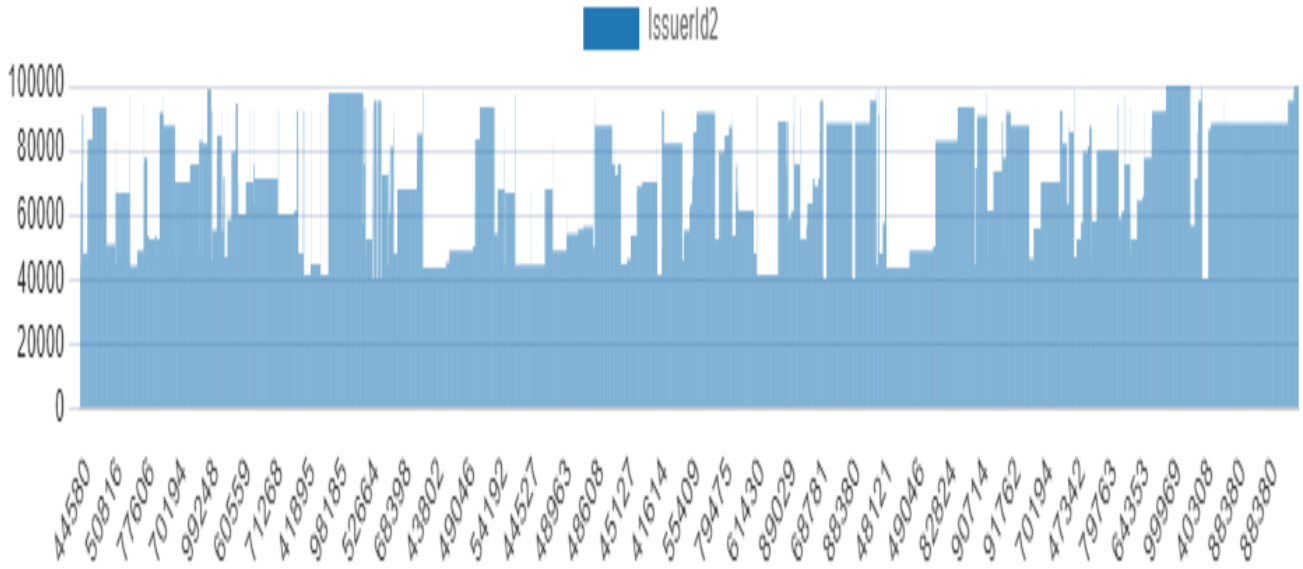


| | A | B | C | D |
|---|---|---|---|---|
| 1 | IssuerId | IssuerId2 | | |
| 2 | 44580 | 44580 | | |
| 3 | 44580 | 44580 | | |
| 4 | 44580 | 44580 | | |
| 5 | 44580 | 44580 | | |
| 6 | 59809 | 59809 | | |
| 7 | 51485 | 51485 | | |
| 8 | 70239 | 70239 | | |
| 9 | 70239 | 70239 | | |
| 10 | 70239 | 70239 | | |
| 11 | 70239 | 70239 | | |
| 12 | 70239 | 70239 | | |
| 13 | 70239 | 70239 | | |
| 14 | 70239 | 70239 | | |
| 15 | 70239 | 70239 | | |
| 16 | 86830 | 86830 | | |
| 17 | 91450 | 91450 | | |
| 18 | 91450 | 91450 | | |

data-1670464600330



| | A | B | C | D |
|---|---|---|---|---|
| 1 | StateCode | count | | |
| 2 | TX | 3146 | | |
| 3 | ND | 251 | | |
| 4 | MS | 462 | | |
| 5 | IN | 1147 | | |
| 6 | MO | 2022 | | |
| 7 | FL | 1305 | | |
| 8 | WI | 1315 | | |
| 9 | NC | 1086 | | |
| 10 | OK | 953 | | |
| 11 | GA | 2241 | | |
| 12 | PA | 3525 | | |
| 13 | AK | 35 | | |
| 14 | WY | 46 | | |
| 15 | LA | 491 | | |
| 16 | TN | 906 | | |
| 17 | NJ | 343 | | |
| 18 | AL | 175 | | |

data-1670464539576