Lab 5, Alexey Rakov, BS19-03, a.rakov@innopolis.university

Ex 1.

Sample Data

Data Output | Explain | Messages | Notifications

	id [PK] integer	name text	address text	review text
99990	99989	Samantha Webb	3432 Dana Lane Suite 014	Daughter foot suddenly again interesting detail instead. Beyond president management employee education ...
99991	99990	Shane Lee	979 Chan Grove	Value brother father process body. Arrive charge his price season where them.
99992	99991	Ashley Wood	771 Carl View Apt. 870	Yourself time stay research base generation station scientist. Friend tend analysis a my. Specific continue poli...
99993	99992	Barry Ferguson	61631 Golden Overpass	Southern lawyer win kid daughter remain. Project success approach social. Wife allow artist realize any issue ...
99994	99993	Robert Williams	7468 Howe Land Suite 965	Add activity church place. Standard reason maybe significant guess be relationship.
99995	99994	Jonathan Oliver	09868 Williams Springs Suite 192	Tax draw stand final far. Hard approach which close. Ability image meeting require.
99996	99995	Lacey Clements	7471 Robert Cape Apt. 926	Central who throughout both list child price. News news herself seek million herself. Check beyond industry fu...
99997	99996	Carrie Kramer	7926 Green Skyway	Receive fall few relationship. Wall brother born.
99998	99997	Susan Carter	33270 Adam Cliffs	Daughter discuss improve. Film TV character her. Center sure despite religious ask. West score ok nice.
99999	99998	Kenneth Harris MD	80397 Hess Stream Suite 603	Sing child message sport. I produce agency thus thought each. Could information child perhaps add way.
100000	99999	Mr. Thomas Farmer	68785 Penny Neck Apt. 037	Will story newspaper beat team today she. Quickly old what sure protect.

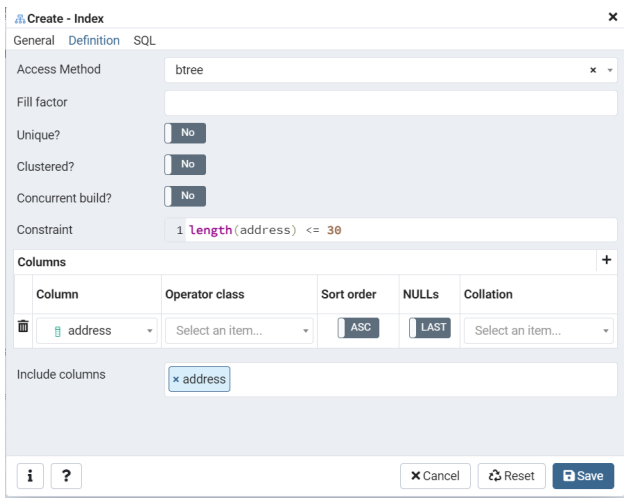
1) First query with analyzing:

```
1  EXPLAIN ANALYZE SELECT * FROM customers_review WHERE length(address) <= 30;
```

Data Output | Explain | Messages | Notifications

	QUERY PLAN text
1	Seq Scan on customers_review (cost=0.00..4534.26 rows=33339 width=211) (actual time=0.014..19.387 rows=3687 loops=1)
2	Filter: (length(address) <= 30)
3	Rows Removed by Filter: 96313
4	Planning Time: 0.047 ms
5	Execution Time: 19.489 ms

Create B-tree Index on Address:



Indexed query with analazying:

Query Editor Query History

```
1  EXPLAIN ANALYZE SELECT * FROM customers_review WHERE length(address) <= 30;
```

Data Output Explain Messages Notifications

	QUERY PLAN text
1	Bitmap Heap Scan on customers_review (cost=167.05..3701.04 rows=33333 width=211) (actual time=0.799..2.480 rows=3687 l...
2	Recheck Cond: (length(address) <= 30)
3	Heap Blocks: exact=2133
4	-> Bitmap Index Scan on address_index (cost=0.00..158.72 rows=33333 width=0) (actual time=0.581..0.581 rows=3687 loops=1)
5	Planning Time: 1.032 ms
6	Execution Time: 2.652 ms

Cost is less a little bit and actual time is less in six time

Python script:

```
import psycopg2

con = psycopg2.connect(database="postgres", user="postgres",
                       password="88664422", host="127.0.0.1", port="5432")

print("Database opened successfully")
cur = con.cursor()
cur.execute("EXPLAIN ANALYZE SELECT * FROM customers_review WHERE length(address) <= 30;")

analyze_fetched = cur.fetchall()
print(analyze_fetched)
```

Result of script (same as pgAdmin query):

```
Database opened successfully
[('Seq Scan on customers_review  (cost=0.00..4534.26 rows=33339 width=211) (actual time=0.022..25.346 rows=3687 loops=1)',),
ilter: 96313',), ('Planning Time: 1.169 ms',), ('Execution Time: 25.453 ms',)]
>>>
```

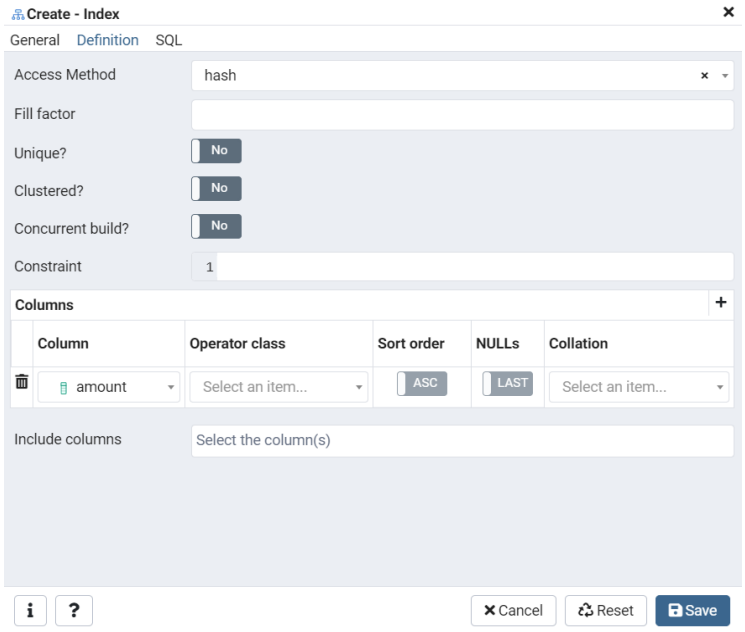
2) Second query with analayzing(chosen table with numbers for more visible results):

```
1  EXPLAIN ANALYZE SELECT * FROM payment WHERE amount = 5.99;
```

Data Output | Explain | Messages | Notifications

	QUERY PLAN text
1	Seq Scan on payment (cost=0.00..290.45 rows=1188 width=26) (actual time=0.017..3.235 rows=1188 loops=1)
2	Filter: (amount = 5.99)
3	Rows Removed by Filter: 13408
4	Planning Time: 0.786 ms
5	Execution Time: 3.297 ms

Create hash index:



Cost is less, actual time is less too

Indexed query with analyzing:

```
1  EXPLAIN ANALYZE SELECT * FROM payment WHERE amount = 5.99;
```

Cost is less, actual time is less too

Data Output | Explain | Messages | Notifications

	QUERY PLAN text
1	Bitmap Heap Scan on payment (cost=41.21..164.06 rows=1188 width=26) (actual time=0.143..0.682 rows=1188 loops=1)
2	Recheck Cond: (amount = 5.99)
3	Heap Blocks: exact=106
4	-> Bitmap Index Scan on payment_index (cost=0.00..40.91 rows=1188 width=0) (actual time=0.120..0.120 rows=1188 loops=1)
5	Index Cond: (amount = 5.99)
6	Planning Time: 1.554 ms
7	Execution Time: 0.759 ms

Python script:

```
import psycopg2

con = psycopg2.connect(database="postgres", user="postgres",
                       password="88664422", host="127.0.0.1", port="5432")

print("Database opened successfully")
cur = con.cursor()
cur.execute("EXPLAIN ANALYZE SELECT * FROM payment WHERE amount = 5.99;")

analyze_fetched = cur.fetchall()
print(analyze_fetched)
```

Script result (same as pgAdmin query without index):

```
Database opened successfully
[('Seq Scan on payment  (cost=0.00..290.45 rows=1188 width=26) (actual time=0.036..3.395
8 loops=1)',), ('  Filter: (amount = 5.99)',), ('  Rows Removed by Filter: 13408',), ('Pl
ime: 3.277 ms',), ('Execution Time: 3.504 ms',)]
>>>
```