Quiz1

Deadline	Friday, 26 February 2021 at 11:59PM
Latest Submission	Monday, 22 February 2021 at 6:19PM
Raw Mark	4.00/4.00 (100.00%)
Late Penalty	N/A
Final Mark	4.00/4.00 (100.00%)

Question 1 (1 mark)

How many times will the read() function be invoked in scanning the file:

```
-rw-r--r-- 1 cs9315 22439 Feb 18 2021 big.txt
```

using the code below

```
char buf[1024];
int nchars;
int in = open("big.txt",O_RDONLY);
while ((nchars = read(in,buf,1024)) > 0) {
    ... do something with contents of buf[] ...
}
close(in);
```

You can assume that the code is run from the directory containing big.txt, that the file is readable, and that all of the appropriate **#include**'s have been done.

```
23
```

✓ Your response was correct.

Mark: 1.00

One read for each 1024-byte block (21 reads); one read for the last block (935 bytes); one read at the end of the file to set nchars to 0 and exit the loop

Question 2 (1 mark)

Consider a relational schema with two tables R(x,y,z) and S(w,x) and an SQL query on this schema:

```
select * from R, S where R.x=S.x and R.y = 2 and S.w > 4;
```

which of the following relational algebra expressions will most likely produce the most efficient evaluation of the query.

Note that Sel[c]R is relation selection, Proj[a,b]R is relational projection, and $(R \ Join \ S)$ is relational join (natural join).

You can assume that 10% of tuples in *R* have an attribute *y* with value 2, and 33% of tuples in *S* have an attribute *w* with value larger than 4.

(a) O	Sel[y=2 and w>4](R Join S)
(b) O	Sel[y=2](Sel[w>4](R Join S))
(c) O	Sel[y=2](R Join (Sel[w>4](S))
(d) O	Sel[w>4]((Sel[y=2](R)) Join S)
(e) ©	(Sel[y=2](R)) Join (Sel[w>4](S))

✓ Your response was correct.

Mark: 1.00

Applying filters (select) before joins typically produces a much more efficient join. And often the filters themselves can make use of indexes.

Question 3 (1 mark)

Under the PGDATA directory are two subdirectories base and global. Under base are a further collection of subdirectories. The subdirectories under base and the global directory itself contain files which hold table data. What is the difference between the tables held under base and those held in global?

(a)	All user tables are located under base ; all catalog tables are located in global .
(b)	All catalog tables are located under base ; all user tables are located under global .
(c)	All user tables and catalog tables are located under base ; global contains log tables.
(d)	User tables and catalog tables are distributed between base and global to minimise file access costs.
(e)	All user tables plus some catalog tables are located under base ; global contains catalog tables shared by all databases.

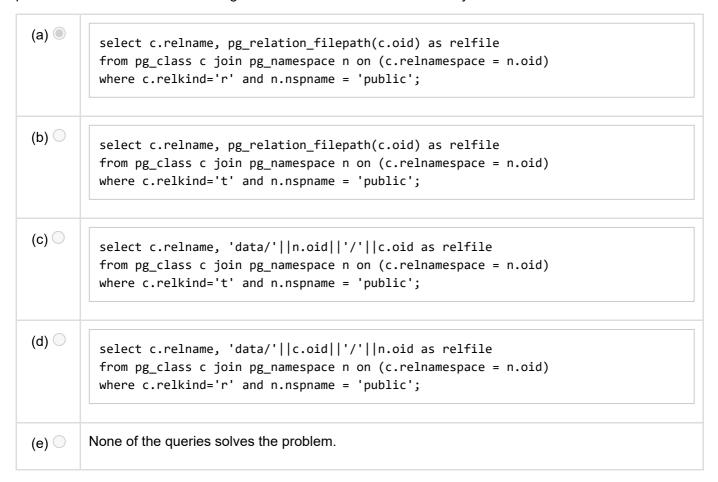
✓ Your response was correct.

Mark: 1.00

The catalog tables under **base** are those containing data local to each database (e.g. pg_class, and any user-defined tables).

Question 4 (1 mark)

Which of the following SQL queries will give a list of the file paths of the data files for all of the tables in the public schema? Names should be given relative to the PGDATA directory.



✓ Your response was correct.

Mark: 1.00

Using pg_relation_filepath() is guaranteed to produce a valid path. Tables have pg_class.relkind='r'.