

19

Wyświetlmy trzy kraje, które w roku 2000 miały dokonane pomiary i na ich podstawie można zaliczyć te kraje jako te o największej liczbie ludności.

Uzupełnij brakujące fragmenty polecenia.

```
select  name ✓ , population ✓ ['2000'] as pop2000
from hbaseCountries
order ✓ by pop2000 desc
limit ✓ 3
```

Pytanie **1**

Poprawnie

Punkty maks.:
4,00

🚩 [Oflaguj pytanie](#)

20

Chcielibyśmy dowiedzieć na jakie trzy pierwsze litery nazw jest najwięcej krajów i jaki jest średni produkt narodowy brutto w tych krajach. Napisz stosowne zapytanie.

Uzupełnij brakujące fragmenty polecenia.

```
select substr( name ✓ ,1,1) as first_letter, count ✓ (*) as how_many, avg ✓ (gdp) as a
from hbaseCountries
group by substr( name ✓ ,1,1)
order by how_many ✓ desc
limit 3 ✓ ;
```

Pytanie **2**

Poprawnie

Punkty maks.:
6,00

🚩 [Oflaguj pytanie](#)

To było proste. Teraz trochę trudniejsze. Chcielibyśmy uzupełnić powyższe zapytanie o sumaryczną liczbę ludności w roku 2000 w tych krajach. Kilka wskazówek:

- Jak wiemy, nie w każdym roku pomiary są dokonywane w każdym kraju – jeśli w roku 2000 nie było pomiarów, bierzemy wcześniejszy najbliższy rok w którym dane o ludności są dostępne.
- Nasza informacja o ludności to mapa, aby sobie z nią poradzić warto skorzystać z funkcji explode. Poniżej przykład użycia funkcji explode. Zwróć uwagę na dwa aliasy. Do czego się one odnoszą? W przypadku jakiego typu danych byłby to jeden alias?

```
SELECT name, year, pop
FROM hbaseCountries LATERAL VIEW explode(population) popTable as year, pop
WHERE name = 'China';
```

Uzupełnij brakujące fragmenty polecenia.

```
SELECT substr( name ✓ ,1,1) as first_letter, count(*) as how_many,
        avg(gdp) as avg_gdp, sum(pop) as sum_pop
FROM ( SELECT name, gdp, year, pop,
        row_number() over (partition by name ✓ order by year ✓ desc) as rn
FROM hbaseCountries LATERAL VIEW explode ✓ (population) popTable as year ✓ , pop
WHERE year <= 2000) tab
WHERE rn = 1
GROUP BY substr( name ✓ ,1,1)
ORDER BY how_many desc
limit 3;
```