

Раздел SQL

Задание 1

Вывести отсортированный по количеству перелетов (по убыванию) и имени (по возрастанию) список пассажиров, совершивших хотя бы 1 полет.

```
SELECT name, COUNT(trip) AS count
FROM Passenger
INNER JOIN Pass_in_trip
ON Passenger.id=Pass_in_trip.passenger
GROUP BY passenger
HAVING COUNT(trip) >= 1
ORDER BY count DESC, name
```

Задание 2

Сколько времени обучающийся будет находиться в школе, учась со 2-го по 4-ый уч. предмет ?

```
SELECT DISTINCT TIMEDIFF(
(SELECT end_pair FROM Timepair WHERE id=4),
(SELECT start_pair FROM Timepair WHERE id=2)
) AS time
FROM Timepair
```

Задание 3

Выведите список комнат, которые были зарезервированы в течение 12 недели 2020 года

```
SELECT DISTINCT Rooms.*
FROM Rooms
JOIN Reservations
ON Rooms.id=Reservations.room_id
WHERE WEEK(start_date, 1) = 12 AND YEAR(start_date)=2020
```

Задание 4

Какой(ие) кабинет(ы) пользуются самым большим спросом?

```
SELECT classroom
FROM Schedule
GROUP BY classroom
HAVING COUNT(classroom) =
(SELECT COUNT(classroom)
FROM Schedule
GROUP BY classroom
ORDER BY COUNT(classroom) DESC LIMIT 1)
```

Задание 6

Составить отчет о битвах кораблей в два суперстолбца.

```
SELECT ROW_NUMBER() OVER(ORDER BY date, name) AS rn,  
NTILE(2) OVER(ORDER BY date, name) AS tile,  
name,  
LAG(date,0,NULL) OVER(ORDER BY date) date_1,  
LEAD(date,3,NULL) OVER(ORDER BY date) date_2  
FROM battles
```

Вашего запроса:					правильного запроса:					
rn	tile	name	date_1	date_2	rn_1	name_1	date_1	rn_2	name_2	date_2
1	1	North Atlantic	1941-05-25 00:00:00.000	1944-10-25 00:00:00.000	1	North Atlantic	1941-05-25 00:00:00.000	4	Surigao Strait	1944-10-25 00:00:00.000
2	1	Guadalcanal	1942-11-15 00:00:00.000	1962-10-20 00:00:00.000	2	Guadalcanal	1942-11-15 00:00:00.000	5	#Cuba62a	1962-10-20 00:00:00.000
3	1	North Cape	1943-12-26 00:00:00.000	1962-10-25 00:00:00.000	3	North Cape	1943-12-26 00:00:00.000	6	#Cuba62b	1962-10-25 00:00:00.000
4	2	Surigao Strait	1944-10-25 00:00:00.000	NULL						
5	2	#Cuba62a	1962-10-20 00:00:00.000	NULL						
6	2	#Cuba62b	1962-10-25 00:00:00.000	NULL						

Ход решения:

Полагаю, что нужно изначально все столбцы разделять через CASE с условиями в виде выражений

Пробовала сделать через

CASE tile

WHEN 1 THEN (SELECT name FROM battles AS name_1)

WHEN 2 THEN (SELECT name FROM battles AS name_2)

Но запрос не сработал и тогда по идее останется еще столбец tile, который выводить не нужно.

То же самое для ROW_NUMBER()

Либо нужно сделать группировку по tile

Раздел математика

Задание 1

В игре «Что? Где? Когда?» в каждом раунде волчок останавливается в секторе номер n , где n равновероятно принимает одно из значений $0, 1, \dots, 13$. При этом играет первый из секторов по часовой стрелке, который ранее не играл. Найдите вероятность того, что после шести раундов сыграют (в любом порядке) секторы $1, 2, \dots, 6$.

Искомое число $\frac{m}{n(14)}$

$n(14)$ – количество вариантов остановки волчка в 14 секторах

$(14)=14^6$ (6 бросаний по условию)

m – число вариантов, при которых выпадут секторы $1, 2, \dots, 6$

Волчок останавливается на секторах 1-6 и никогда на остальных. Значит вероятность выпадения секторов 1-6 равна вероятности невыпадения остальных секторов. Допустим рассматривается 7 секторов и вероятность выпадения 1-6 и невыпадения оставшегося.

Тогда

$n(7)$ – количество вариантов

$(7)=7^6$

$$\frac{1}{7} = \frac{m}{7^6}$$

$m=7^5$

В таком случае искомая вероятность

$$\frac{7^5}{14^6} = \frac{1}{448}$$

Ответ: $\frac{1}{448}$

Задание 2

Аналитик рынка ценных бумаг оценивает среднюю доходность определенного вида акций. Случайная выборка из 16 дней показала, что средняя доходность по акциям данного типа составляет 8% с выборочным средним квадратическим отклонением в 4%. Предполагая, что доходность акции подчиняется нормальному закону распределения, определите 99%-ый доверительный интервал для средней доходности интересующего аналитика вида акций.

Условие:

$n=16$

$X_{\text{среднее}}=8$ по выборке

$\sigma=4$

уровень доверия $(1-\alpha)=0,99$

$\alpha=0,01$

Решение:

- Так как интервал двусторонний:

$$\frac{\alpha}{2} = 0,05$$
$$1 - \frac{\alpha}{2} = 0,995$$

- Найдено по таблице для нормального распределения:

$$Z_{1-\frac{\alpha}{2}}=2,58$$

$$P(X_{\text{ср}} - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq X_{\text{ср}} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

Где μ – среднее в генеральной совокупности

Нижняя граница интервала:

$$X_{\text{ср}} - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 8 - 2,58 \cdot \frac{4}{\sqrt{16}} = 5,42$$

Верхняя граница интервала:

$$X_{\text{ср}} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 8 + 2,58 \cdot \frac{4}{\sqrt{16}} = 10,58$$

Ответ: С вероятностью 0,99 $\mu \in [5,42; 10,58]$

Задание 3

Мужчины и женщины по-разному оценивают положительные человеческие качества. Предложили мужчинам и женщинам на основе десятибалльной шкалы (10 баллов – это максимум) оценить важность следующих пяти качеств в представителях противоположного пола:

Качества	Ум	Доброта	Красота	Юмор	Работоспособность
Мужчины	7	8	8	5	7
Женщины	10	5	3	8	10

Найдите тесноту связи между этими данными, рассматривая данные, как выборочные наблюдения случайных величин. Сделайте вывод о том, насколько близки или далеки мужчины и женщины в оценках качеств партнеров.

Решение:

Тесноту связи между имеющимися данными можно определить с помощью коэффициента корреляции. Расчет по задаче проводила в Excel через формулу =КОРРЕЛ(массив1; массив2), где массив1 это мужчины и массив2 женщины (если поменять их местами, результат будет тот же).

Расчетная формула:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{j=1}^n (Y_j - \bar{Y})^2}},$$

Таким образом

$$r = -0,5243225$$

Знак минус означает, что одна переменная уменьшается по мере того, как увеличивается другая (обратная зависимость).

Значение близко к -0,5, что данные имеют слабую обратную взаимосвязь. Однако это не означает наличие причинно-следственной связи между рассматриваемыми данными.

Раздел Основы программирования

Дан массив целых чисел `nums` и целое число `target`. Написать функцию, возвращающую индексы элементов, дающих в сумме число `target`?

Python:

```
import random
import numpy
nums = []
nums = [random.randint(1,10) for i in range(1,20)]
print('Массив: %s' % nums)
if type(nums) == int:
    target = 10
    for i in range(nums+1):
        target += i
print('Индексы: %s' % nums.index())
```

Ход решения:

- 1) Задаем массив из 20 любых чисел в интервале от 1 до 10
- 2) При условии, что массив `nums` состоит из целых чисел, присваиваем значение переменной `target`
- 3) В цикле массива суммируются попарно все числа
- 4) Выводятся индексы элементов, суммы которых равны `target`