

Время выполнения задания – 180 мин., язык – русский и английский.

1. Дайте развернутые ответы (на английском языке):

1.1. What is a "dark data" and how they can be used for analysis? (15 баллов)

1.2. How big data analytics affect the development of smart device? (15 баллов)

2. Решите задачи:

2.1. (15 баллов)

Компания, занимающаяся маркетинговыми исследованиями, ведет базу данных для поддержания и автоматизации информационных потоков. Сведения о заказах включают следующие данные: номер, название заказа, даты старта и окончания выполнения заказа, перечень услуг, название контрагента, ФИО контактного лица, телефон, стоимость заказа, номер договора, дата договора, статус заказа.

Перечень оказываемых услуг содержит каталог из видов услуг и перечня маркетинговых исследований, их краткого описания. Для каждого заказа определяются сотрудники, которые будут выполнять работу, причем база данных содержит справочник со специализацией каждого сотрудника по видам оказываемых услуг и проводимым маркетинговым исследованиям.

Данные по контрагенту содержат: ИНН, вид контрагента, наименование, полное наименование, страну регистрации, КПП, ОГРН. По одному договору могут выполняться несколько заказов.

Требуется:

а) используя любую общепринятую нотацию, нарисовать схему базы данных, удовлетворяющую третьей нормальной форме, указать типы и направления связи;

б) сделать подробное описание таблиц с расшифровкой имен полей, указанием типов и свойств данных, ключевых полей;

в) используя операторы языка SQL, написать запросы для получения следующей информации:

- количество заказов, начатых выполняться в каждом месяце за текущий год. Результат представить в виде «месяц-количество заказов» по возрастанию месяцев;

- номера договоров, стоимость заказов по которым в прошлом месяце превышает среднюю стоимость заказов по договорам за прошлый месяц.

2.2. (10 баллов)

Найти значение параметра α , при котором функция $\alpha(1 + x^2)^{-2}$ является плотностью распределения вероятностей некоторой случайной величины при $x \in (-\infty, +\infty)$.

2.3. (15 баллов)

Некоторая случайная величина характеризуется биномиальным распределением, задаваемым формулой Бернулли $P_n(m) = C_n^m p^m (1-p)^{n-m}$. Какое распределение будет характерно для данной случайной величины при $n \rightarrow \infty$?

2.4. (15 баллов)

Обозначения и термины.

Алфавитом называется конечное множество символов.

Непустой цепочкой в алфавите A называется произвольная последовательность $d[1], \dots, d[n]$, где $n \geq 1$, и для $k = 1, \dots, n$ элемент $d[k]$ - это символ в алфавите A (принято не указывать запятую в обозначениях таких последовательностей).

Таким образом, $bbbddcccc$ – это непустая цепочка в алфавите $\{b, d, c\}$.

Выражения вида $AxB \rightarrow AyB$, где A, B - цепочки символов (возможно, пустые), x - это символ, y - символ или цепочка символов, будем понимать как контекстные подстановки, когда x заменяется на y , если слева от x стоит A , а справа от x стоит B .

Постановка задачи. Пусть $T = \{x, y, z\}$, назовем это множество основным алфавитом.

Пусть $N = \{A, B, C, D, E\}$, назовем это множество вспомогательным алфавитом.

Пусть P - следующая система контекстных подстановок:

$$S \rightarrow xAB, S \rightarrow xB, A \rightarrow xAC,$$

$$A \rightarrow xC, B \rightarrow Dz, D \rightarrow y,$$

$$CD \rightarrow CE, CE \rightarrow DE, DE \rightarrow DC,$$

$$Cz \rightarrow Dzz.$$

Какое множество цепочек в алфавите T можно получить всевозможными подстановками из P (как угодно много подстановок), если всегда первой применяется какая-либо подстановка с левой частью S ? **Ответ обосновать.**

2.5. (15 баллов)

Плотность распределения вероятностей с носителем $(-\infty, +\infty)$ непрерывной случайной величины имеет следующий вид:

$$\rho(x) = A[1 - (1 - q)x^2]^{1/(1-q)},$$

где $q < 3$ – параметр распределения, A – нормировочная константа.

Найти математическое ожидание случайной величины.

