Olympiad for students and graduates – 2017
Demonstrative version and guidelines for the area
«Business Informatics»

Profile: «Big Data Systems»

You have 180 min to complete this task.

1. Give detailed answers

1.1. What means data driven approach to enterprise business process management?
1.2. Describe the possibilities of modern business intelligence systems

2. Solve a task

2.1. The company that owns a network of data processing centers (DPC) stores in a relational database the following data: name of the DPC, URL, contact information, telephone.

Range of services of each DPC is determined by its profile and contains a list of services with the tariffs, which is located on the site of each DPC. Each data center keeps records of the services provided for each day where it stores the name of service, quantity (units), current sale price for its service and details of counterparties.

Each order is assigned a unique number, and the contracting party one order can have multiple different services from the list of DPC.

In the provision of services and the establishment of tariffs the company focused on world prices for telecommunications services and the prices of its partners and suppliers. This is the base for company's actual cost calculation for each services, including the settlement date.

Do the following:

1. draw the database schema that meets the third normal form, with primary and foreign keys, type and direction of relation, using any accepted notation
2. make a detailed description of the tables with a breakdown of the field names, data types and properties
3. write queries to retrieve the following information using MySQL:
   - the order numbers in which last month the revenue from the sale of services exceeds the average value of orders for the previous month, indicate the amount of revenue; sort order numbers in descending order of amounts of revenue;
   - the number of orders per month, from the beginning of the calendar year to the present for each DPC.

2.2. Coefficients p and q of the quadratic equation $x^2+px+q=0$ are independent random values with uniform distribution at the interval [0,a] (a>0).
What is the probability that the roots of this equation are real numbers?

2.3. Denotations and terms.
Let *d* be an arbitrary symbol, for instance, a digit or a letter of Latin alphabet. Then *d* to the power 1 (designated as d ^ 1) is another denotation of the symbol *d.*
If *n* is an arbitrary integer, and *n > 1*, then d to the power *n* (designated as d_1 ^ n) is the sequence consisting of *n* occurrences of the symbol *d*. For instance, *b^3 d^2 c^4* is the string *bbbddcccc*.

Let the following set (in other words, language) *L* be given:
L = {b^3m c^k d^2n, где m, k, n >=1}.

Let *Help-set* be an arbitrary finite set of symbols not containing the Latin letters *b, c, d*.
The symbols from the set *Help-set* will be considered as auxiliary symbols.

Let's say that an arbitrary finite set of expressions of the form x --> y is a productions system for the set *Help-set* if and only if x is an arbitrary symbol from the set *Help-set*, and the expression *y* satisfies one of the following conditions:

(1) *y* is one of the letters *b, c, d;*
(2) *y* is an expression of the form *h E*, where *h* is one of the letters *b, c, d,* and *E* is any symbol from the set *Help-set*;
(3) *y* is the expression *void-string*.

The application of the substitution of the form x --> *void-string* means the deletion of the symbol *x*.

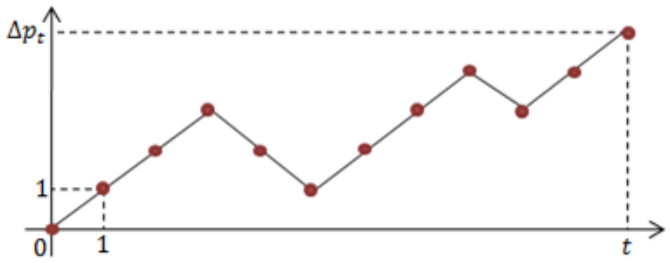
Task statement. Invent such set of auxiliary symbols *Help-set* including the symbol *start* and such productions system *P* for the set *Help-set* that the given set (or language) *L* is the collection of all expressions (or strings) in the alphabet *{b, c, d}* that can be obtained by any substitutions from *P* (arbitrary many substitutions) on condition that at the first step any substitution with the left side *start* is applied.

2.4. One of the provisions of the market chaos theory is that the behavior of prices of market is described by a random walk.

The simplest model of price dynamics is a one-dimensional model of a discrete random walk (random process with discrete time).

Consider time series $\{\Delta p_t\}$ of price variations Δp with discrete time t. Let's assume that initially $\Delta p_0=0$, and sequentially at each time t =1,2,3,… price variation equally probable to increase or reduce by 1.

Please, estimate the number of possible time series with positive Δp (see Graph), with initial value $\Delta p_0=0$ and final value Δp_t ($0<\Delta p_t\leq t$).



2.5. Assume, that the rate of change of number of invoices per day that were included by transport company employee into database is proportional to the number of invoices per day that are not included to database.

An employee of transport company has to include 200 invoices per day to the database.

A new employee after the job placement in 1 week included 50 invoices per day.

Estimate the number of invoices that this employee could put into the database per day after one more week.