

Варианты к заданиям лабораторной работы 6

Задание 3

№ варианта	Описание набора данных
1	<p>Фрагмент набора данных, полученных по результатам опроса учащихся португальских школ.</p> <p><i>Age</i> – возраст учащегося,</p> <p><i>Reason</i> – основная причина выбора данной школы,</p> <p><i>Studytime</i> – время, потраченное на домашние задания в течение недели,</p> <p><i>Failures</i> – количество несданных дисциплин,</p> <p><i>G2</i> – итоговая оценка по математике за год (от 0 до 20 баллов).</p>
2	<p>Фрагмент набора данных с информацией о сотрудниках компании.</p> <p><i>Age</i> – возраст сотрудника,</p> <p><i>Department</i> – сфера деятельности,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Length_of_service</i> – стаж работы в организации,</p> <p><i>Previous_year_rating</i> – рейтинг сотрудника за предыдущий год по пятибальной шкале.</p>
3	<p>Фрагмент набора данных с результатами опроса об удовлетворенности пассажиров авиакомпании совершенным перелетом.</p> <p><i>Age</i> – возраст пассажира,</p> <p><i>Gender</i> – пол пассажира,</p> <p><i>Class</i> – класс обслуживания в самолете,</p> <p><i>Departure Delay in Minutes</i> – задержка вылета (в минутах),</p> <p><i>Departure/Arrival time convenient</i> – уровень удовлетворенности временем отправления/прибытия (от 1 – не удовлетворен до 5 – полностью удовлетворен).</p>
4	<p>Фрагмент набора данных с результатами Бостонского марафона 2019 года.</p> <p><i>Age</i> – возраст бегуна,</p> <p><i>Gender</i> – пол бегуна,</p> <p><i>Country</i> – гражданство,</p> <p><i>Result_sec</i> – результат забега (в секундах)</p> <p><i>Rank_Gender</i> – место в рейтинге по результатам забега среди женщин/мужчин.</p>
5	<p>Фрагмент набора данных с информацией о госпитализированных больных с диабетом в больницах США в период с 1999 по 2008 годы.</p> <p><i>Gender</i> – пол пациента,</p> <p><i>Age</i> – возрастной диапазон,</p> <p><i>Weight</i> – вес пациента (в фунтах),</p> <p><i>Number_emergency</i> – количество вызовов скорой помощи к пациенту за год, предшествующий текущей госпитализации,</p> <p><i>Glyburide</i> – назначение/изменение дозировки препарата Глибурид (<i>up</i> – дозировка увеличена, <i>steady</i> – дозировка не изменилась, <i>up</i> – дозировка понижена, <i>no</i> – препарат не назначался).</p>

№ варианта	Описание набора данных
6	<p>Фрагмент набора данных с информацией о зарплатах специалистов в области Data Science.</p> <p><i>Work_year</i> – за какой год указана сумма заработной платы,</p> <p><i>Salary_in_usd</i> – сумма заработной платы в долларах,</p> <p><i>Remote_ratio</i> – коэффициент удаленной работы (0 – работа в офисе, объем работы, выполняемой удаленно, составляет менее 20% от общего объема), 50 – гибридный режим, 100 – удаленная работа (объем работы, выполняемой удаленно, составляет более 80% от общего объема),</p> <p><i>Company_location</i> – страна, в которой расположена компания-работодатель,</p> <p><i>Company_size</i> – размер компании.</p>
7	<p>Фрагмент набора данных с информацией о клиентах банка.</p> <p><i>Age</i> – возраст клиента,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Geography</i> – страна проживания,</p> <p><i>EstimatedSalary</i> – расчетная заработная плата,</p> <p><i>Tenure</i> – сколько лет является клиентом банка.</p>
8	<p>Фрагмент набора данных с результатами обследования женщин с раком молочной железы.</p> <p><i>Age</i> – возраст пациентки,</p> <p><i>Race</i> – раса пациентки,</p> <p><i>Differentiate</i> – степень дифференцировки опухоли,</p> <p><i>Grade</i> – стадия рака,</p> <p><i>Regional Node Examined</i> – количество обследованных регионарных лимфоузлов.</p>
9	<p>Набор данных с характеристиками звезд.</p> <p><i>Temperature (K)</i> – температура поверхности (в кельвинах),</p> <p><i>Radius(R/R_o)</i> – радиус звезды, рассчитанный относительно радиуса Солнца,</p> <p><i>Star color</i> – цвет звезды по результатам спектрального анализа,</p> <p><i>Spectral Class</i> – класс звезды по спектральной классификации звёзд,</p> <p><i>Star type</i> – тип звезды (0 – коричневый карлик, 1 – красный карлик, 2 – белый карлик, 3 - звезда главной последовательности, 4 – сверхгигант, 6 – гипергигант).</p>
10	<p>Фрагмент набора данных с информацией о клиентах, их потребительском поведении и результатах рекламных компаний.</p> <p><i>Year_Birth</i> – год рождения,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Marital_Status</i> – семейное положение,</p> <p><i>Kidhome</i> – количество несовершеннолетних детей,</p> <p><i>Income</i> – годовой доход семьи клиента.</p>

№ варианта	Описание набора данных
11	Фрагмент набора данных с информацией о физическом состоянии и образе жизни людей из латиноамериканских стран. <i>Age</i> – возраст анкетированного, <i>Height</i> – рост (м), <i>CALC</i> – ответ на вопрос «Как часто Вы употребляете алкоголь?», <i>CH2O</i> – ответ на вопрос «Сколько примерно литров воды Вы выпиваете за день?», <i>MTRANS</i> – ответ на вопрос «Каким видом транспорта Вы обычно пользуетесь?».
12	Фрагмент набора данных с информацией о вызовах, поступивших в пожарные части Торонто в период 2011 по 2018 годы. <i>Extent_Of_Fire</i> – оценка масштаба пожара, <i>Civilian_Casualties</i> – число жертв среди гражданского населения, <i>Count_of_Persons_Rescued</i> – число спасенных людей, <i>Estimated_Dollar_Loss</i> – предполагаемая сумма убытков (в долларах), <i>Business_Impact</i> – оценка сроков восстановления работоспособности фирмы.
13	Фрагмент набора данных с информацией о сотрудниках организации. <i>Age</i> – возраст сотрудника, <i>Gender</i> – пол сотрудника, <i>Education</i> – уровень образования, <i>JoiningYear</i> – год приема на работу в организацию, <i>PaymentTier</i> – уровень заработной платы (более высокое значение соответствует более высокой заработной плате).
14	Фрагмент набора данных с информацией о треках, входивших в топ-100 песен на Spotify в период с 2010 по 2019 года. <i>Genre</i> – жанр трека, <i>Year_released</i> – год выпуска, <i>DB</i> – громкость (в децибелах), <i>Duration</i> – продолжительность трека (в секундах), <i>Artist_type</i> – число исполнителей.
15	Фрагмент набора данных из исследования по выбору вида эвкалипта для борьбы с эрозией почвы. <i>Sp</i> – вид эвкалипта, <i>Year</i> – год высадки саженцев, <i>Ht</i> – высота деревьев (м), <i>Utility</i> – оценка влияния деревьев на состояние почвы, <i>Rainfall</i> – количество осадков (мм в год).
16	Фрагмент набора данных, полученных по результатам опроса учащихся португальских школ. <i>Age</i> – возраст учащегося, <i>Medu</i> – уровень образования матери, <i>Mjob</i> – сфера работы матери, <i>Health</i> – оценка текущего состояния здоровья (от 1 – очень плохое состояние до 5 – очень хорошее состояние), <i>G2</i> – итоговая оценка по математике за год (от 0 до 20 баллов).

№ варианта	Описание набора данных
17	<p>Фрагмент набора данных с информацией о сотрудниках компании.</p> <p><i>Age</i> – возраст сотрудника,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Recruitment_channel</i> – источник найма сотрудника,</p> <p><i>No_of_trainings</i> – количество пройденных обучений за прошедший год,</p> <p><i>Avg_training_score</i> – средний балл по результатам пройденных обучений.</p>
18	<p>Фрагмент набора данных с результатами опроса об удовлетворенности пассажиров авиакомпании совершенным перелетом.</p> <p><i>Type of Travel</i> – тип поездки пассажира,</p> <p><i>Class</i> – класс обслуживания в самолете,</p> <p><i>Flight Distance</i> – дальность полета (км),</p> <p><i>Arrival Delay in Minutes</i> – задержка прибытия (в минутах),</p> <p><i>On-board service</i> – уровень удовлетворенности бортовым обслуживанием (от 1 – не удовлетворен до 5 – полностью удовлетворен).</p>
19	<p>Фрагмент набора данных, предоставленных при заполнении онлайн-заявки на кредит.</p> <p><i>Property_Area</i> – тип места проживания,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Dependents</i> – число иждивенцев,</p> <p><i>LoanAmount</i> – сумма кредита (тысяч долларов),</p> <p><i>Loan_Amount_Term</i> – срок кредитования (месяцев).</p>
20	<p>Фрагмент набора данных с информацией о госпитализированных больных с диабетом в больницах США в период с 1999 по 2008 годы.</p> <p><i>Race</i> – раса пациента,</p> <p><i>Age</i> – возраст пациента,</p> <p><i>Weight</i> – диапазон веса пациента в фунтах,</p> <p><i>Time_in_hospital</i> – количество дней между госпитализацией и выпиской,</p> <p><i>Glipizide</i> – назначение/изменение дозировки препарата Глипизид (<i>up</i> – дозировка увеличена, <i>steady</i> – дозировка не изменилась, <i>up</i> – дозировка понижена, <i>no</i> – препарат не назначался).</p>
21	<p>Фрагмент набора данных с информацией о зарплатах специалистов в области Data Science.</p> <p><i>Designation</i> – название профессии,</p> <p><i>Experience</i> – уровень специалиста (EN - джун, MI - мидл, SE - сеньор, EX – руководитель/директор)</p> <p><i>Working_Year</i> – за какой год указана сумма заработной платы,</p> <p><i>Salary_In_Rupees</i> – сумма заработной платы в рупиях,</p> <p><i>Remote_Working_Ratio</i> – коэффициент удаленной работы (0 – работа в офисе, объем работы, выполняемой удаленно, составляет менее 20% от общего объема), 50 – гибридный режим, 100 – удаленная работа (объем работы, выполняемой удаленно, составляет более 80% от общего объема).</p>

№ варианта	Описание набора данных
22	<p>Фрагмент набора данных с информацией о клиентах банка.</p> <p><i>Gender</i> – пол клиента,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Balance</i> – сумма на счете,</p> <p><i>CreditScore</i> – баллы кредитного рейтинга,</p> <p><i>Tenure</i> – сколько лет является клиентом банка.</p>
23	<p>Фрагмент набора данных с результатами обследования женщин с раком молочной железы.</p> <p><i>Marital Status</i> – семейное положение пациентки,</p> <p><i>Tumor Size</i> – размер опухоли (в мм),</p> <p><i>Differentiate</i> – степень дифференцировки опухоли,</p> <p><i>Grade</i> – стадия рака,</p> <p><i>Regional Node Positive</i> – количество регионарных лимфоузлов с обнаруженными опухолевыми клетками.</p>
24	<p>Набор данных с характеристиками звезд.</p> <p><i>Temperature (K)</i> – температура поверхности (в кельвинах),</p> <p><i>Absolute magnitude(Mv)</i> – абсолютная визуальная звездная величина,</p> <p><i>Star color</i> – цвет звезды по результатам спектрального анализа,</p> <p><i>Spectral Class</i> – класс звезды по спектральной классификации звёзд,</p> <p><i>Star type</i> – тип звезды (0 – коричневый карлик, 1 – красный карлик, 2 – белый карлик, 3 - звезда главной последовательности, 4 – сверхгигант, 6 – гипергигант).</p>
25	<p>Фрагмент набора данных с информацией о клиентах, их потребительском поведении и результатах рекламных компаний.</p> <p><i>Year_Birth</i> – год рождения,</p> <p><i>Education</i> – уровень образования,</p> <p><i>Marital_Status</i> – семейное положение,</p> <p><i>NumDealsPurchases</i> – количество покупок, совершенных со скидкой,</p> <p><i>MntWines</i> – сумма, потраченная на вино за последние 2 года.</p>
26	<p>Фрагмент набора данных с информацией о физическом состоянии и образе жизни людей из латиноамериканских стран.</p> <p><i>Age</i> – возраст,</p> <p><i>Height</i> – рост (м),</p> <p><i>CAEC</i> – ответ на вопрос «Вы делаете перекусы между основными приемами пищи?»</p> <p><i>FAF</i> – ответ на вопрос «Сколько дней в неделю Вы занимаетесь физической активностью?»</p> <p><i>MTRANS</i> – ответ на вопрос «Каким видом транспорта Вы обычно пользуетесь?».</p>

№ варианта	Описание набора данных
27	<p>Фрагмент набора данных с информацией о вызовах, поступивших в пожарные части Торонто в период 2011 по 2018 годы.</p> <p><i>Incident_Station_Area</i> – номер пожарной части, к зоне ответственности которой относится адрес, по которому поступил вызов,</p> <p><i>Status_of_Fire_On_Arrival</i> – оценка степени пожара по прибытии,</p> <p><i>Incident_Station_Area</i> – площадь возгорания,</p> <p><i>Possible_Cause</i> – предполагаемая причина возникновения пожара,</p> <p><i>Estimated_Dollar_Loss</i> – предполагаемая сумма убытков (в долларах),</p> <p><i>Civilian_Casualties</i> – число жертв среди гражданского населения.</p>
28	<p>Фрагмент набора данных с информацией о сотрудниках организации.</p> <p><i>Age</i> – возраст сотрудника,</p> <p><i>Gender</i> – пол сотрудника,</p> <p><i>Education</i> – уровень образования,</p> <p><i>JoiningYear</i> – год приема на работу в организацию,</p> <p><i>ExperienceInCurrentDomain</i> – стаж сотрудника в его текущей сфере деятельности.</p>
29	<p>Фрагмент набора данных с информацией о треках, входивших в топы песен на Spotify в 2022 году.</p> <p><i>Artist_names</i> – имя/название исполнителя,</p> <p><i>Peak_rank</i> – самое высокое место в рейтинге, до которого доходил трек,</p> <p><i>Weeks_on_chart</i> – количество недель, который трек продержался в топе,</p> <p><i>Duration</i> – продолжительность трека (в секундах),</p> <p><i>Mode</i> – лад (0 – минор, 1 – мажор).</p>
30	<p>Фрагмент набора данных из исследования по выбору вида эвкалипта для борьбы с эрозией почвы.</p> <p><i>Sp</i> – вид эвкалипта,</p> <p><i>Year</i> – год высадки саженцев,</p> <p><i>DBH</i> – диаметр ствола дерева, измеренный на высоте груди человека (1,3 м выше уровня земли),</p> <p><i>Utility</i> – оценка влияния деревьев на состояние почвы,</p> <p><i>Frost</i> – количество осадков (мм в год).</p>

Задание 4

№ варианта	Среднее квадратическое отклонение признака X
1	0,5
2	5,5
3	8
4	8,5
5	7,5
6	1
7	0,5
8	6
9	3
10	4
11	12
12	12
13	4,5
14	12
15	2
16	8,5
17	6
18	10
19	3
20	8
21	2,5
22	3
23	2
24	11,5
25	1
26	0,5
27	10,5
28	5
29	9,5
30	11,5