

Пакет задач

Задача 4

Вариант 1

Имеются данные (Time Almanac 2000) о гестационном периоде и ожидаемой продолжительности жизни 10 видов животных, представленные в таблице.

Гестационный период (в днях) и ожидаемая продолжительность жизни (лет) для 10 видов животных

Animal	Gestation Period	Life Expectancy (yrs)
Cat	63	11
Chicken	22	7,5
Dog	63	11
Duck	28	10
Goat	151	12
Lion	108	10
Parakeet	18	8
Pig	115	10
Rabbit	31	7
Squirrel	44	9

1. Определите источник корреляционной зависимости и выделите независимую и зависимую переменную.
2. Постройте уравнение парной линейной зависимости, оцените его качество с помощью основных показателей.
3. Сделайте выводы.

Вариант 2

Исследователи умственных способностей человека хотят знать, влияет ли объем мозга на умственные способности. Для этого коллективом авторов (Willirman and others, 1991) было проведено тестирование выборочной совокупности на IQ и проведена магнитно-резонансная диагностика для определения веса мозга (MRI), результаты представлены количеством ненулевых пикселей по 18 проекциям. Результаты исследований приведены в таблице.

Результаты сканирования и тестов на IQ выборки студентов психологического факультета, имеющих оценки выше среднего

Gender	MRI Count	IQ
Female	816932	133
Female	951545	137
Female	991305	138
Female	833868	132
Female	856472	140
Female	852244	132
Female	790619	135
Female	866662	130
Female	857782	133
Female	948066	133
Male	949395	140
Male	1001121	140
Male	1038437	139
Male	965353	133
Male	955466	133
Male	1079549	141
Male	924059	135
Male	955003	139
Male	935494	141
Male	949589	144

1. Постройте линейное уравнение регрессии для IQ в качестве зависимой переменной для всех студентов, затем для мужчин и женщин отдельно.
2. Оцените полученные модели.
3. Представьте себе, что преподаватель хочет спрогнозировать IQ для новенькой студентки Алисы, чей MRI равен 830 000. Выберите нужное уравнение и выполните прогноз.
4. Сделайте выводы.

Вариант 3

Используя таблицу задачи 2 варианта, выполните следующие задания:

1. Постройте таблицу сопряженности по признакам «пол» и «IQ», разделив совокупность на 3 группы по IQ.
2. Протестируйте гипотезу о связи этих двух признаков.
3. Рассчитайте коэффициенты взаимной сопряженности.
4. Сделайте вывод.

Вариант 4

Специалисты по страхованию автомобилей проанализировали данные о выданных правах на вождение автомобиля и авариях на транспорте.

Данные статистической службы о выданных правах и авариях на транспорте по полу и возрастным группам

Age Group	Number of Male Licensed Drivers	Number of Crashes Involving a Male	Number of Female Licensed Drive	Number of Crashes Involving a Female
16	816	244	764	178
17	1198	233	1115	175
18	1342	243	1212	164
19	1454	229	1333	145
20 - 24	7866	951	7394	618
25 - 29	9356	899	8946	595
30 - 34	10121	875	9871	571
35 - 39	10521	901	10439	566
40 - 44	9776	692	9752	455
45 - 49	8754	667	8710	390
50 - 54	6840	390	6763	247
55 - 59	5341	290	5258	165
60 - 64	4565	218	4486	133
65 - 69	4234	191	4231	121
70 - 74	3604	167	3749	104
75 - 79	2563	118	2716	77
80 - 84	1400	61	1516	45
> 85	767	34	767	20

1. По данной таблице оцените возможность регрессионного моделирования между следующими парами факторов: возрастом и числом выданных прав мужчинам, возрастом и числом выданных прав женщинам.
2. Выберите признаки (пары), наиболее тесно связанные друг с другом и постройте уравнения с помощью ANOVA. Оцените их качество.
3. Сделайте прогноз числа аварий для мужчин и женщин при 8700 прав, выданных мужчинам и 8500 выданных женщинам.

Вариант 5

Проанализируйте кейс *Physical Strength and Job Performance* (адрес кейса он-лайн http://onlinestatbook.com/case_studies_rvls/physical_strength/index.html) и составьте маршрутную карту исследования, отмечая на каждом этапе наиболее важные результаты и принятое решение. Ответьте на вопросы, которые вы найдете на страницах кейса.