**Задача.** Изучается зависимость средней ожидаемой продолжительности жизни от нескольких факторов.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Страна | y | x1 | x2 | x3 | x4 | x5 | x6 |
| Мозамбик | 47+a | 3.0 | 2.6 | 2.4 | 113+b | 1.4+a | 8.8 |
| Ангола | 47+a | 4.9 | 3.1 | 2.8 | 124+b | 5.3+a | 1.2 |
| Чад | 48+a | 2.6 | 2.5 | 2.5 | 117+b | 1.2+a | 0.2 |
| Бурунди | 49+a | 2.3 | 2.6 | 2.7 | 98+b | 8.0+a | 9.1 |
| Буркина-фасо | 49+a | 2.9 | 2.8 | 2.1 | 99+b | 3.6+a | 0.1 |
| Мали | 50+a | 2.0 | 2.9 | 2.7 | 123+b | 4.7+a | 1.9 |
| Бенин | 50+a | 6.5 | 2.9 | 2.5 | 95+b | 4.8+a | 0.4 |
| Мавритания | 51+a | 5.7 | 2.5 | 2.7 | 96+b | 7.7+a | 6.2 |
| Конго | 51+a | 7.6 | 2.9 | 2.6 | 90+b | 9.7+a | 3.7 |
| Мадагаскар | 52+a | 2.4 | 3.1 | 3.1 | 89+b | 7.7+a | 3.8 |
| Нигерия | 53+a | 4.5 | 2.9 | 2.8 | 80+b | 6.3+a | 5.5 |
| Непал | 55+a | 4.3 | 2.5 | 2.4 | 91+b | 3.2+a | 4.2 |
| Того | 56+a | 4.2 | 3.0 | 2.8 | 88+b | 8.6+a | 2.0 |
| Гаити | 57+a | 3.4 | 2.0 | 1.7 | 72+b | 3.6+a | 1.6 |
| Зимбабве | 57+a | 7.5 | 2.4 | 2.2 | 55+b | 8.9+a | 9.6 |
| Камерун | 57+a | 7.8 | 2.9 | 3.1 | 56+b | 9.3+a | 9.2 |
| Папуа-новая гвинея | 57+a | 9.0 | 2.3 | 2.3 | 64+b | 3.9+a | 6.5 |
| Бангладеш | 58+a | 5.1 | 1.6 | 2.1 | 79+b | 1.9+a | 9.9 |
| Кения | 58+a | 5.1 | 2.7 | 2.7 | 58+b | 7.2+a | 1.2 |
| Гана | 59+a | 7.4 | 2.8 | 2.7 | 73+b | 5.7+a | 6.3 |
| Пакистан | 60+a | 8.3 | 2.9 | 3.3 | 90+b | 9.1+a | 6.3 |
| Индия | 62+a | 5.2 | 1.8 | 2.0 | 68+b | 5.5+a | 8.3 |
| Египет | 63+a | 14.2 | 2.0 | 2.7 | 56+b | 5.5+a | 4.9 |
| Индонезия | 64+a | 14.1 | 1.6 | 2.5 | 51+b | 0.1+a | 3.1 |
| Юар | 64+a | 18.6 | 2.2 | 2.4 | 50+b | 5.2+a | 0.5 |
| Марокко | 65+a | 12.4 | 2.0 | 2.6 | 55+b | 9.4+a | 0.2 |
| Филиппины | 66+a | 10.6 | 2.2 | 2.7 | 39+b | 7.2+a | 1.9 |
| Гватемала | 66+a | 12.4 | 2.9 | 3.5 | 44+b | 6.4+a | 9.6 |
| Перу | 66+a | 14.0 | 2.0 | 3.1 | 47+b | 7.1+a | 8.1 |
| Гондурас | 67+a | 7.0 | 3.0 | 3.8 | 45+b | 1.3+a | 0.7 |
| Республика эль-сальвадор | 67+a | 9.7 | 2.2 | 3.4 | 36+b | 1.1+a | 9.8 |
| Бразилия | 67+a | 20.0 | 1.5 | 1.6 | 44+b | 7.9+a | 6.1 |
| Турция | 67+a | 20.7 | 1.7 | 2.1 | 48+b | 6.5+a | 1.8 |
| Никарагуа | 68+a | 7.4 | 3.1 | 4.0 | 46+b | 6.8+a | 9.2 |
| Парагвай | 68+a | 13.5 | 2.7 | 2.9 | 41+b | 10.0+a | 7.0 |
| Китай | 69+a | 10.8 | 1.1 | 1.1 | 34+b | 2.4+a | 0.5 |
| Эквадор | 69+a | 15.6 | 2.2 | 3.2 | 36+b | 3.8+a | 8.1 |
| Тунис | 69+a | 18.5 | 1.9 | 3.0 | 39+b | 5.4+a | 0.3 |
| Таиланд | 69+a | 28.0 | 0.9 | 1.3 | 35+b | 2.7+a | 3.6 |
| Белоруссия | 70+a | 15.6 | 0.2 | 0.2 | 13+b | 6.8+a | 0.1 |
| Алжир | 70+a | 19.6 | 2.2 | 4.1 | 34+b | 9.1+a | 5.1 |
| Польша | 70+a | 20.0 | 0.3 | 0.6 | 14+b | 8.8+a | 3.1 |
| Доминиканская республика | 71+a | 14.3 | 1.9 | 2.6 | 37+b | 5.0+a | 5.9 |
| Венесуэла | 71+a | 29.3 | 2.3 | 3.0 | 23+b | 3.3+a | 5.2 |
| Малайзия | 71+a | 33.4 | 2.4 | 2.7 | 12+b | 2.8+a | 2.6 |
| Мавритания | 71+a | 49.0 | 1.3 | 1.8 | 16+b | 1.9+a | 4.8 |
| Шри-ланка | 72+a | 12.1 | 1.3 | 2.0 | 16+b | 6.9+a | 8.7 |
| Словакия | 72+a | 13.4 | 0.3 | 0.7 | 11+b | 4.8+a | 9.8 |
| Мексика | 72+a | 23.7 | 1.9 | 2.8 | 33+b | 0.2+a | 5.4 |
| Тринидад | 72+a | 31.9 | 0.8 | 1.8 | 13+b | 7.3+a | 5.7 |
| Чили | 72+a | 35.3 | 1.5 | 2.1 | 12+b | 7.1+a | 4.5 |
| Республика корея | 72+a | 42.4 | 0.9 | 1.9 | 10+b | 1.9+a | 7.2 |
| Панама | 73+a | 22.2 | 1.7 | 2.4 | 23+b | 7.7+a | 3.9 |
| Уругвай | 73+a | 24.6 | 0.6 | 1.0 | 18+b | 7.2+a | 6.7 |
| Аргентина | 73+a | 30.8 | 1.3 | 2.0 | 22+b | 8.8+a | 2.6 |
| Ямайка | 74+a | 13.1 | 1.0 | 1.8 | 13+b | 2.5+a | 4.6 |
| Дания | 75+a | 78.7 | 0.3 | 0.1 | 6+b | 3.2+a | 5.1 |
| Новая зеландия | 76+a | 60.6 | 1.4 | 1.5 | 7+b | 5.2+a | 2.5 |
| Финляндия | 76+a | 65.8 | 0.5 | 0.1 | 5+b | 6.9+a | 0.0 |
| Сингапур | 76+a | 84.4 | 2.0 | 1.7 | 4+b | 2.7+a | 6.2 |
| Испания | 77+a | 53.8 | 0.2 | 1.0 | 7+b | 1.4+a | 5.8 |
| Ирландия | 77+a | 58.1 | 0.5 | 1.7 | 6+b | 7.2+a | 8.5 |
| Израиль | 77+a | 61.1 | 3.5 | 3.5 | 8+b | 8.1+a | 9.6 |
| Австралия | 77+a | 70.2 | 1.1 | 1.4 | 6+b | 4.7+a | 0.2 |
| Австрия | 77+a | 78.8 | 0.8 | 0.5 | 6+b | 7.5+a | 4.3 |
| Бельгия | 77+a | 80.3 | 0.4 | 0.5 | 8+b | 0.5+a | 6.6 |
| США | 77+a | 100.0 | 1.0 | 1.1 | 8+b | 2.8+a | 2.5 |
| Греция | 78+a | 43.4 | 0.6 | 0.9 | 8+b | 9.8+a | 3.4 |
| Франция | 78+a | 70.0 | 0.5 | 0.8 | 6+b | 2.3+a | 3.0 |
| Италия | 78+a | 73.7 | 0.2 | 0.4 | 7+b | 7.8+a | 1.2 |
| Нидерланды | 78+a | 73.9 | 0.7 | 0.6 | 6+b | 4.7+a | 2.3 |
| Канада | 78+a | 78.3 | 1.3 | 1.0 | 6+b | 1.3+a | 2.6 |
| Швейцария | 78+a | 95.9 | 1.0 | 0.8 | 6+b | 4.6+a | 3.6 |
| Швеция | 79+a | 68.7 | 0.6 | 0.3 | 4+b | 9.8+a | 6.9 |
| Гонконг | 79+a | 85.1 | 1.6 | 1.3 | 5+b | 7.1+a | 9.7 |
| Япония | 80+a | 82.0 | 0.3 | 0.6 | 4+b | 1.4+a | 10 |

a – последняя цифра номера зачетной книжки;

b – предпоследняя цифра номера зачетной книжки.

В таблице приняты обозначения у – средняя ожидаемая продолжительность жизни, лет; х1 – ВВП в паритетах покупательной способности; х2 – темпы прироста населения по сравнению с предыдущим годом, %; х3 – темпы прироста рабочей силы по сравнению с предыдущим годом, %; х4 – коэффициент младенческой смертности, %; x5, x6 – дополнительные параметры.

1. Необходимо с использованием библиотек Python:

1) с использованием библиотеки Yellowbrick сделать попарное сравнение переменных корреляцией Пирсона, результаты вывести на график;

2) аналогичный график, «тепловую карту», вывести с использованием библиотеки seaborn;

3) стандартизировать переменные у, x1, x2, x3, x4, x5, x6;

4) для стандартизированных переменных оценить коэффициенты и рассчитать модельные значения для линейной регрессионной модели,

у=b0+b1\*x1+ b2\*x2+ b3\*x3+ b4\*x4+ b5\*x5+ b6\*x6;

5) отобрать значимые признаки и оценить коэффициенты линейной регрессионной модели, рассчитать модельные значения;

6) рассчитать модельные значения для модели К-ближайшие соседи;

7) рассчитать модельные значения для модели Дерево решений;

8) определить коэффициент детерминации для всех моделей, сравнить, выбрать лучшую модель;

9) вывести графики остатков, сделать заключение о качестве модели;

10) подготовить подробный отчёт о выполненной работе.