

1. Что такое разведочный анализ данных?
2. Какова главная задача разведочного анализа данных?
3. Как называется этап обработки данных, включающий удаление пропусков и выбросов?
4. Какие данные используются в разведочном анализе?
5. Как называются точки данных, сильно отличающиеся от остальных?
6. Что означает высокий коэффициент вариации?
7. Что измеряет коэффициент Пирсона?
8. Как оценивается качество регрессионной модели?
9. Что включает в себя EDA (Exploratory Data Analysis)?
10. Зачем очищаются данные перед началом анализа?
11. Что такое стандартное отклонение?
12. Чем отличается дисперсия от стандартного отклонения?
13. Чем характеризуется равномерное распределение?
14. Что называют масштабированием данных?
15. Как определяется межквартильный диапазон (IQR)?
16. Пусть имеются две переменные: $X=\{1, 2, 3\}$, $Y=\{4, 5, 6\}$. Посчитайте коэффициент корреляции Пирсона $r(X,Y)$.
17. Чему равен коэффициент вариации, если стандартное отклонение равно 5, а среднее значение — 20?
18. Если $Q_1=10$, $Q_2=20$, $Q_3=30$, какой будет межквартильный диапазон IQR?
19. При каком уровне доверительного интервала в статистике принимается предположение о высокой достоверности результата?
20. Дан набор данных: $\{10, 15, 20, 25, 30\}$. Рассчитайте среднеквадратичное отклонение.