1. **Что такое корреляция?**

Корреляция - это статистическая мера, которая описывает степень взаимосвязи или взаимосвязанности между двумя переменными. Она показывает, насколько сильно и в каком направлении две переменные изменяются вместе. Корреляция может быть положительной, отрицательной или равной нулю, что указывает на направление и силу взаимосвязи между переменными.

1. **Что показывает коэффициент корреляции? Каким он может быть?**

Коэффициент корреляции показывает степень линейной зависимости между двумя переменными. Он может принимать значения от -1 до 1. Коэффициент корреляции +1 означает положительную линейную корреляцию, -1 означает отрицательную линейную корреляцию, а 0 означает отсутствие линейной корреляции. Чем ближе значение к 1 или -1, тем сильнее линейная взаимосвязь между переменными.

1. **Что представляет из себя модель простой линейной регрессии?**

Модель простой линейной регрессии представляет собой статистическую модель, которая описывает линейную зависимость между одним независимым (предикторным) и одним зависимым (целевым) переменными. Она представляется уравнением прямой линии, которая наилучшим образом соответствует данным. Модель простой линейной регрессии обычно выражается уравнением y = mx + b, где y - зависимая переменная, x - независимая переменная, m - коэффициент наклона (наклон прямой), а b - свободный член (точка пересечения с осью y).

1. **Чем множественная линейная регрессия отличается от простой линейной регрессии?**

В отличие от простой линейной регрессии, где есть только один предиктор (независимая переменная), множественная линейная регрессия включает два или более предиктора для прогнозирования зависимой переменной. Таким образом, уравнение множественной линейной регрессии имеет вид y = b0 + b1x1 + b2x2 + ... + bnxn, где y - зависимая переменная, x1, x2, ..., xn - независимые переменные (предикторы), b0 - свободный член, b1, b2, ..., bn - коэффициенты регрессии.

1. **Что характеризует коэффициент детерминации?**

Коэффициент детерминации (R^2) представляет собой статистическую меру, которая показывает, как много изменчивости зависимой переменной объясняется моделью регрессии. Он измеряет пропорцию дисперсии зависимой переменной, которая может быть объяснена независимыми переменными модели. R^2 может принимать значения от 0 до 1, где 0 означает, что модель не объясняет вариацию зависимой переменной, а 1 означает, что модель объясняет всю вариацию зависимой переменной.

1. **Как сложность модели влияет на коэффициент детерминации?**

Чем сложнее модель, тем выше может быть коэффициент детерминации. Это связано с тем, что более сложные модели могут лучше подстраиваться под данные и объяснять больше вариации. Однако использование слишком сложных моделей может привести к переобучению, когда модель хорошо работает на обучающих данных, но плохо обобщается на новые данные. Поэтому важно найти баланс между сложностью модели и ее способностью обобщать данные.

**Описание модели:**

Модель простой линейной регрессии позволяет нам предсказывать значение целевой переменной (пробег на галлоне) на основе одного признака (вес автомобиля). В нашем случае признаком является вес автомобиля в фунтах (lbs).

Описание модели:

* **Целевая переменная** **(y)**: Пробег на галлоне (mpg) - это количественная переменная, которую мы пытаемся предсказать.
* **Признаки (X):** Вес автомобиля (lbs) - это наш единственный признак, на основе которого мы строим предсказания.
* **Уравнение модели**: mpg = intercept + weight \* coef\_weight, где intercept - это коэффициент пересечения (intercept), а coef\_weight - это коэффициент веса (weight).
* **Коэффициенты модели**: Мы обучили модель и оценили коэффициенты intercept и coef\_weight, которые позволяют нам предсказывать пробег на галлоне на основе веса автомобиля

Оценка качества модели: Мы оценили качество модели на обучающем наборе данных и получили значение коэффициента детерминации (R^2), который показывает, как много дисперсии в целевой переменной объясняется нашей моделью. В нашем случае модель с одним параметром (весом автомобиля) показала приемлемую точность в предсказании пробега на галлоне.

Таким образом, наша модель простой линейной регрессии может быть использована для прогнозирования пробега на галлоне на основе веса автомобиля, что может быть полезным при анализе эффективности и экономичности автомобилей.