

EDA assignment Question #1

①

$$m(a+bX) = a + b \times m(x)$$

$$= \frac{1}{n} \sum_{i=1}^n (a + bx_i)$$

$$= \frac{1}{n} \left(\sum_{i=1}^n a + \sum_{i=1}^n bx_i \right)$$

$$= \frac{1}{n} \left(an + b \sum_{i=1}^n x_i \right)$$

$$= a + b \cdot \frac{1}{n} \sum_{i=1}^n x_i$$

$$= a + b(m(x))$$

Factored out b

②

$$\text{cov}(X, a+bY) = b \times \text{cov}(X, Y)$$

$$= \frac{1}{N} \sum_{i=1}^N (x_i - m(x)) (a + by_i - m(a + by))$$

$$= \frac{1}{N} \sum_{i=1}^N (x_i - m(x)) (a + by_i - a - bmy)$$

$$= \frac{1}{N} \sum_{i=1}^N (x_i - m(x)) (by_i - bmy)$$

$$= b \underbrace{\frac{1}{N} \sum_{i=1}^N (x_i - m(x)) (y_i - m(y))}_{\text{cov}(X, Y)}$$

$$= b \times \text{cov}(X, Y)$$

$$(3) \text{Cov}(a+bX, a+bX) = b^2 \text{Cov}(X, X)$$

$$= b^2 \cdot \frac{1}{N} \sum_{i=1}^N (x_i - m(x))(x_i - m(x))$$

$$= b^2 \cdot \frac{1}{N} \sum_{i=1}^N (x_i - m(x))^2 \rightarrow = s^2$$

$$= b \left(\frac{1}{N} \sum_{i=1}^N (x_i - m(x)) \right)^2 \cdot b \left(\frac{1}{N} \sum_{i=1}^N (x_i - m(x)) \right)^2$$

$$= \text{Cov}(a+bX, a+bX)$$

4) Yes it is still a median because a non-decreasing transformation does not affect order in data, which is what mean is based on.

This does not apply to IQR or range, however, because unlike median it is actually based on the distance between 2 points

5) yes if $g()$ is only increasing, but not necessarily true if otherwise