EDA hw

Q1:

1) show that $m(a+bX) = a + b \times m(X)$

$m(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \rightarrow m(a+bX) = \frac{1}{n}\sum_{i=1}^{n}(a+bX_i)$

$= m(a+bX) = \frac{1}{n}\sum_{i=1}^{n} a + \frac{1}{n}\sum_{i=1}^{n} bX_i$

because a is a constant: $m(a+bX) = \frac{na}{n} + \frac{b}{n}\sum_{i=1}^{n} X_i$

therefore: $m(a+bX) = a + b \times m(X)$

2) show that $cov(X, a+bY) = b \times cov(X,Y)$

$cov(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))(Y_i - m(Y))$

↳ $cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))((a+bY_i) - m(a+bY))$

$m(a+bY) = \frac{1}{n}\sum_{i=1}^{n}(a+bY_i) = a + b \cdot m(Y)$

↳ substitute: $cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))((a+bY_i) - (a+b\,m(Y))$

↳ a cancels: $cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))(b(Y_i - m(X)))$

↳ factor out b: $cov(X, a+bY) = b \cdot \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))(Y_i - m(Y))$

↳ simplify: $cov(X, a+bY) = b \times cov(X,Y)$

3) show that $cov(a+bX, a+bX) = b^2 cov(X,X)$ w/ $cov(X,X) = S^2$

setting $X=Y$ for $cov(X,Y) \rightarrow cov(X,X) = \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))^2 = S^2$

$cov(a+bX, a+bX) = \frac{1}{N}\sum_{i=1}^{N}[(a+bX_i) - m(a+bX)] \cdot [(a+bX_i) - m(a+bX)]$

$m(a+bX) = \frac{1}{n}\sum_{i=1}^{n}(a+bX_i) = a + b\,m(X)$

rewrite: $cov(a+bX, a+bX) = \frac{1}{N}\sum_{i=1}^{N}(bX_i - b\,m(X))(bX_i - b\,m(X))$

factor b: $cov(a+bX, a+bX) = b^2 \cdot \frac{1}{N}\sum_{i=1}^{N}(X_i - m(X))^2$

simplify: $cov(a+bX, a+bX) = b^2 \cdot cov(X,X)$

$= b^2 \cdot S^2$

4) -No, a non-decreasing transformation of the median is not always the median of the transformed variable. In strictly increasing functions, the order of the data points is is preserved and so transforming X doesn't change the number of X's below and above the median. However, in non-decreasing but not strictly increasing functions, this is not necessarily true.

- Yes, this applies to any quartile. Distortions of the data in any quartile can occur after transformation.

- Because the IQR changes according to g(x), it can still transform at the quartile level.

- For the range, a transformation may change the range and make it smaller or bigger

5) No. A non-decreasing function can have flat areas, where g(x) is constant, m(X) can change into different values. and then the median of g(x) ≠ the g(x) value of the median X.