Part 4 in the assignment file

Q1 EDA

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

$m(X) = \frac{1}{N}\sum_{i=1}^{N} x_i$

$m(a+bX) = \frac{1}{N}\sum_{i=1}^{N}(a+b\cdot x_i)$

$\sum_{i=1}^{N} a = N\cdot a$

$m(a+bX) = \frac{1}{N}\left(\sum_{i=1}^{N} a + \sum_{i=1}^{N} b\cdot x_i\right)$

$\sum_{i=1}^{N} b\cdot x_i = b\cdot \sum_{i=1}^{N} x_i$

$m(a+bX) = \frac{1}{N}(N\cdot a + b\cdot \sum_{i=1}^{N} x_i)$

$m(a+bX) = a + b\cdot \frac{1}{N}\sum_{i=1}^{N} x_i$

$m(a+bX) = a + b\cdot m(X)$

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

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

$cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))((a+b\cdot y_i) - m(a+bY))$

$m(a+bY) = a + b\cdot m(Y)$

$cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))((a+b\cdot y_i) - (a+b\cdot m(Y)))$

$cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))(b\cdot(y_i - m(Y)))$

$cov(X, a+bY) = b\cdot \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))(y_i - m(Y))$

$cov(X, a+bY) = b\cdot cov(X,Y)$

3) Show that $cov(a+bX, a+bX) = b^2 cov(X,X)$, and $cov(X,X) = S^2$

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

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

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

$\frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))^2 = cov(X,X) = S^2$

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

5) No, it is not always true that $m(g(X)) = g(m(X))$. The sample mean $m(X)$ is a linear operator, but $g(X)$ may not be linear.
For example, if $g(X) = X^2$, then: $m(g(X)) = m(X^2) \neq m(X))^2 = g(m(X))$
This equality only holds if g is a linear function, such as $g(X) = a+bX$
For non-linear transformations, $m(g(X)) \neq g(m(X))$

Scanned with CamScanner