a. Mean subtraction :
First search the mean from all data, and then subtract each data with the mean.
Normalization :
After the data already subtracted with the mean, divide that data with standard deviation of all data, then the resulting result is normalized.

b. Result of Xavier and He initialization, optimizer adam; activation : relu

Model	Xavier	He
No hidden layer	83.04;82.98;83.36=83.13	83.67;83.82;83.52=83.67
1 Conv2D+Max Pooling	90.13;90.5;90.4=90.34	90.12;90.03;89.81=89.99

He initialization should be better than Xavier if we are using relu activation

c. First Model : No hidden layer, just straight from input layer (28x28) to output layer (10). The accuracy are around 83% with using adam as optimizer.
Second Model : 1 Convolutional layer with 32 filter and with a 3x3 kernel. The accuracy is around 90% with using adam as optimizer and relu as activation.
By adding a convolutional layer, the accuracy increased by 7 point percentage. This means that having an additional layer will help to get better feature representation.

d. Result of Xavier and He initialization, optimizer rmsprop; activation: relu

Model	Xavier	He
No hidden layer	83.98;82.4;83.34=83.24	82.48;81.6;83.52=82.5
1 Conv2D+Max Pooling	90.19;90.25;90.02=90.15	90.11;90.27;89.6=89.99

Overall Adam optimizer will give better result than RMSprop. Because Adam not only store exponentially decaying average of past squared gradient like RMSprop, Adam also keeps an exponentially decaying average of past gradient.

e. Result of Xavier and He initialization, optimizer adam; activation : selu

Model	Xavier	He
No hidden layer		
1 Conv2D+Max Pooling	89.57;89.3;89.64=89.50	88.65;87.99;89.1=88.58

Relu gives better results than selu, and selu is more computationally expensive than relu.

f. Result of Xavier and He initialization, optimizer adam; activation : relu

Model	Xavier	He
1 Conv2D+Max Pooling (L2 Regularizer)	90.53;90.43;90.64=90.53	90.27;89.85;90.28=90.13
1 Conv2D+Max Pooling (Dropout 0.3)	90.13;90.62;90.63=90.46	90.23;90.27;89.51=90.00

By using regularizer will also improve the accuracy results.