**Problem statement**

The manager of the XYZ mall has approached us with this data where he has tried gathering some details regarding the customers who visit the mall. The given dataset is a simple spreadsheet where you can see columns like customer ID, gender of the customer, age of the customer, the annual income of the customer given to us in thousand dollars and the spending score which ranges from 1 to 100. The gender, age, and annual income was gathered through a feedback survey and the spending score was calculated by the manager on the basis of several factors like how often does the customer visit the mall, how much of amount does he spend on shopping and also on entertainment and food. Based on this, from the manager’s perspective, a good spending score which is close to 100 indicates a good customer and spending score close to 1 indicates a bad customer.

Now the manager has come up to us with this data saying that he has an offer in his mind which he wants to roll out to specific customers by personally calling them. Since the data in the real world is huge, this cannot be done manually, since the manager cannot call each and every customer. So what we can do to find out the potential customers for the manager so that he can call only those specific customers and roll out the offer to them.

You are required to work upon this data and find out the optimum number of segments that can be generated, segment the customers and help the manager with the potential customers.