**Project on Amazon Kindle Books**

We have Amazon kindle data is available for around 60k books. It has

1. Tittle
2. Author
3. Price
4. Money saved
5. Pages
6. Description
7. Publisher
8. Customer reviews
9. Rating

Data has been provided in the file named “*Amazon Kindle.csv*”. Answer-

1. Who is the best author?
2. Which are the top 10 bestselling books? Based on
   1. Number of copies
   2. Dollar value/Revenue
3. Is rating dependent on any of the attributes? If yes, on which attributes?
4. Cluster the books based on the content it has. (Based on description)
5. Which book genre/topic has highest rating? – If not able to solve, even the approach will do.

Please provide the R/python code and document the sequence of steps you performed during the process along with the explanation of the output *ppt* format.

**All the Best!**