**Cloud Computing for Data Analysis**

**VIDEO CASE 05 : Finding Similar Items**

Watch following videos:

**Video 1:** <https://youtu.be/wrkVnwaKTjo>

**Video 2:** <https://youtu.be/ubqGFxHeg7Q>

Video 1 gives a sample example for Jaccard coefficient and its limitations

Video 2 gives a sample example for Cosine similarity

**Video Case Questions:**

**D1:** The sky is blue

**D2:** The sun in bright

**Query:** The sun in the sky is bright

1. Find Jaccard coefficient for the above documents (D1 and D2) for the query Q
2. What is the advantage of using cosine similarity over Jaccard coefficient?
3. Where do you think, these measures can be used?