Nama : Ricky arianto  
Nim : 2022010  
Prodi : Teknik Komputer

Bagian 1:

1. Similarity
2. Untuk mengekstrak kata kata
3. Karena web dan penyimpanan data besar membutuhkan infrastruktur komputasi yang menyimpan dan memproses data tersebut
4. Map reduce framework

Bagian 2:

1. 