

1. Jawaban pertanyaan sebagai berikut,
 - a. Kata yang memiliki jumlah karakter terbanyak adalah dipenjarakannya dengan jumlah karakter sebanyak 15
 - b. Jumlah kalimat yang terdapat dalam korpus adalah 54263
 - c. Jumlah kata unik dalam korpus adalah 22848
 - d. Rata – rata jumlah kalimat untuk setiap dokumen adalah
 - e. Jumlah kata unik yang memiliki jumlah karakter kelipatan lima adalah 9816
 - f. Distribusi kata dalam korpus
 - g. Banyak kata yang dapat saya temukan dan diubah adalah 1758
 - h. Jumlah kemungkinan acara yang saya temukan adalah 0
2. Regular expression sebagai berikut,
 - a. Token yang terdiri dari tiga angka berurutan yang jika dibaca dari depan atau belakang (atau terbalik) menunjukkan angka yang sama (contoh : 212)
 - b. Semua angka kelipatan 7 (contoh : 7, 14, 21, dst.)
 - c. Semua angka dengan karakter romawi (contoh : IV)