

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

گزارش کار بخش برنامه نویسی R تمرین سری چهارم

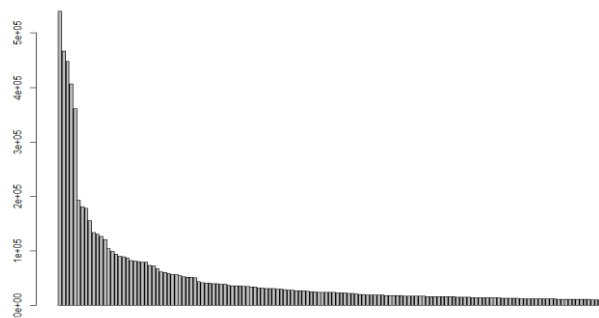
فائزه قربانپور ۹۳۱۰۳۴۳۶

۱. تصادفی حرف ننیم!

با توجه به داده های موجود در فایل txt برخی کلمات خیلی زیاد استفاده می شوند، و بسیاری از کلمات به مقدار کم و این موجب می شود هیستوگرام مقدار تکرار این داده ها به شکل دو خط عمود برهم در ربع اول مختصات در نقطه (0,0) باشد بنابراین در رسم نمودار این داده ها تکرار های کمتر از ۱۰۰۰ در نظر گرفته نشده است.



نمودار ۱: کلماتی که تکرارشان بیشتر از ۱۰۰۰ بار است.



نمودار ۲: کلماتی که تکرارشان کمتر از ۱۰۰۰۰ بار است.

با ایجاد متن تصادفی و شمردن میزان تکرار کلمات به نمودار زیر می‌رسیم.



نمودار ۳: کلمات تصادفی ایجاد شده که تکرارشان بیشتر از ۱۰ بار است.

در هر دو نمودار برخی کلمات بسیار بیشتر از بقیه تکرار شده‌اند ولی در قسمت اول این مقدار در حدود ۱۰۰۰۰۰ می‌باشد ولی در متن تصادفی ما در حدود ۶۰۰ است که این نشان می‌دهد ما تصادفی صحبت نمی‌کنیم و از عمد برخی کلمات را بسیار بسیار بیشتر از بقیه به کار می‌بریم.

ولی در صورتی که تعداد تکرار کلمات متن تصادفی خود را زیاد و زیادتر کنیم با تقریب مناسبی شاید به صحبت کردن ما نزدیک‌تر بشوند.

۲. روح های ناآرام

همانطور که به ازای دلتاهای متفاوت در نمودار ها دیده می شود، تفاوت زیادی بین مقدارهای دو نامساوی نیست در اصول ریاضی نیز نامساوی چپی شف از مارکوف به دست میآید ، در نتیجه این دو نامساوی معادل هم اند.

