در این برنامه یک فایل DOCUMENT.TXT به عنوان ورودی گرفته می شود که شامل چندین ایمیل است و هر ایمیل با تگ باز <BODY> و تگ بسته </BODY> در این فایل قرار می گیرد و در نتیجه ی این برنامه ایمیل هایی که هرزنامه تشخیص داده شده است را به کاربر نشان می دهد البته اسنادی که نشان داده می شود به ترتیب میزان ویژگی هرزنامه بودنشان است به عبارت دیگر ویژگی هرزنامه در ایمیل های اول بیشتر از ایمیل های دیگر است که نشان داده شده است.( از بالا به پایین این ویژگی کاهش میابد.)

نمونه فایل ورودی sample\_input.txt

همانند این فایل ایمیل های خود را در DOCUMENT.txt قرار دید.

<BODY>

…….

</BODY>

1. فایل DOCUMENT.txt خالی است و شما می توانید ایمیل ها را در آن کپی کرده البته فراموش نشود که تگ باز و بسته هر ایمیل را قرار دهید.
2. از آن جایی که ممکن حجم فایل ورودی زیاد باشد برای این منظور که با سرعت بیشتری برنامه اجرا شود ابتدا input\_query() را کامنت کرده و دو تابع write\_start\_unmergefile() و main\_dictionary() را اجرا کرده که فایل های مورد نیاز برای برنامه ساخته شود.
3. سپس برنامه اماده اجرا شدن است و در این قسمت دو تابع که در مرحله قبل ذکر کردم کامنت کرده و تابع سوم را فقط اجرا کرده و تعداد ایمیل هایی که می خواهید مشاهده کنید که هرزنامه است را وارد کنید تا به شما ایمیل ها را نشان دهد.:)