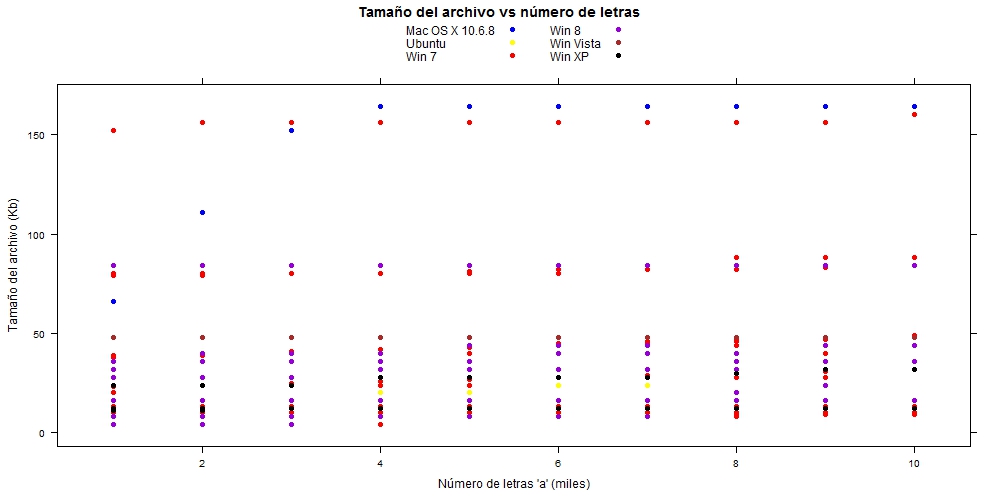
Análisis Variable Sistema Operativo:

En primera instancia vamos a analizar que sucede con la base de datos unificada, cuando hacemos un diagrama de dispersión y agrupamos por sistema operativo, en un primer acercamiento obviaremos los distintos procesadores, formatos etc.

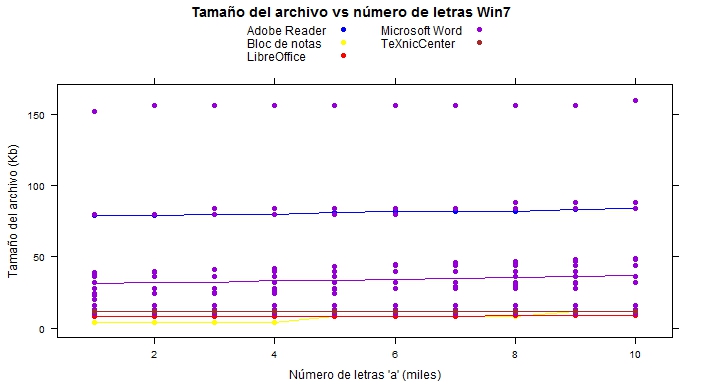
La grafica es la siguiente:

*# HACER GRAFICA DE LINEAS TAMBIEN ¡! , donde cada línea sea un SO? .*

Se puede observar que Win XP, Ubuntu y Win Vista siempre se mantienen relativamente abajo, sin muchas fluctuaciones en los puntos de las diferentes muestras, Mac al ser una sola muestra no es muy significativa, sin embargo vemos que Win 7 en mayor medida y Win 8 en menor tienen en algunas observaciones para un mismo número de letras un muy bajo tamaño de archivo y para otras con el mismo número de letras un muy alto tamaño de archivo , por lo tanto se propone estudiar con mas detalles estos dos Sistemas operativos, muy posiblemente tenga que ver con el procesador de texto y el formato de archivo.

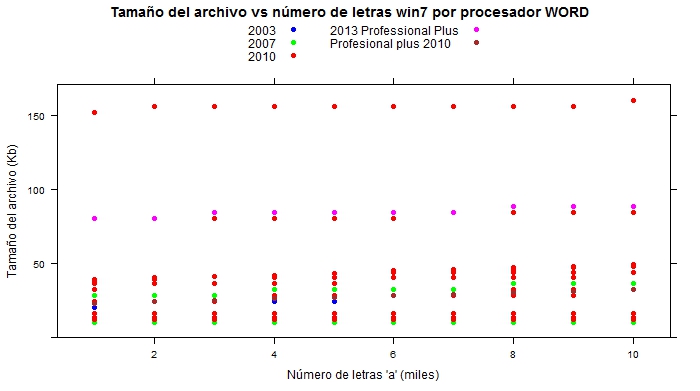
Paso siguiente se tomo una muestra de solamente de los SO win7 y se procedió a graficarlos agrupándolos por el procesador de texto.

Se puede apreciar que adobe reader , Libre office, block de notas , Texnicenter, son bajos y relativamente constantes mientras que Microsoft Word al ser el más abundante se encuentra en casi todo el diagrama , muy disperso, por lo cual toca adentrarse más aun y ver a qué se debe se debe las variaciones que produce .



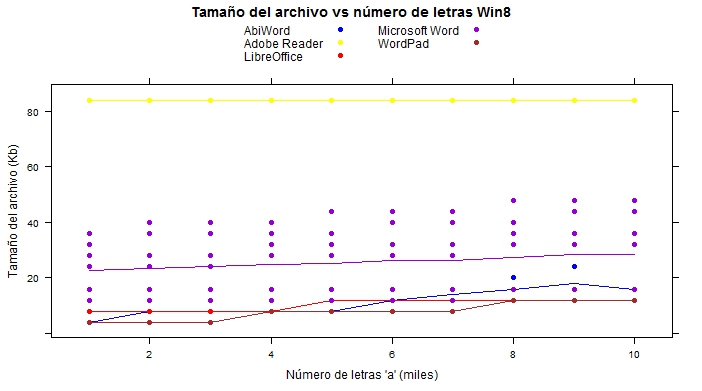
Como ya se menciono los datos más dispersos otra vez son los de Microsoft Word por lo cual se propone aislarlos ya analizar las variables que podrían afectarlo.

Se procede entonces a aislar aun más el procesador y centrarse en el procesador Microsoft Word de Windows 7.

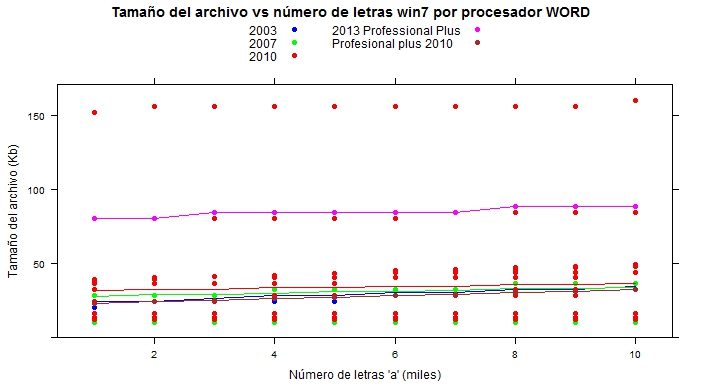


Entre las distintas versiones del procesador Word se propondría ver qué sucede con los archivos de Word 2010, hacer el análisis de las distintos formatos de los archivos, este análisis se vera mas adelante.

Se analizara ahora que sucede con Windows 8 (Win 8)



Se puede observar que los datos más dispersos otra vez son los de Microsoft Word por lo cual se propone aislarlos ya analizar las variables que podrían afectarlo.

Se procede entonces a aislar aun más el procesador y centrarse en el procesador Microsoft Word de Windows 8.

Entre las distintas versiones del procesador Word se propondría ver qué sucede con los archivos de Word 2010 son los más dispersos , se propone hacer el análisis de las distintos formatos de los archivos de Word 2010 .

**2. CRITERIOS DE DEPURACION:**

Se encontraron varias observaciones y se unificaran de la siguiente manera.

Variable escritorio:

|  |  |
| --- | --- |
| escritorio | Escritorio |
| Escritorio | Portátil |
| portatil |
| Portatil |
| portátil |

Variable procesador:

|  |  |
| --- | --- |
| word | Microsoft Word |
| Microsoft Office Word |
| Micorsoft word |
| Microsoft Office Word |
| Word |
| word |
| microsoft word |
| Microsof Word |
| LibreOffice Writer | LibreOffice |
| PDF | Adobe Reader |
| Adobe Reader |

Variable sistema operativo:

|  |  |
| --- | --- |
| windows 7 | Win 7 |
| windows 7 home basic |
| windows 7 professional |
| windows 7 ultimate |
| Windows 7 Home Basic |
| Windows 7 Professional |
| Windows 7 Ultimate |
| wimdows 7 |
| Windows 7 ultimate |
| Windows 7 Home basic |
| Window7 |
| Windows 7 |
| Windows7 |

|  |  |
| --- | --- |
| windows xp | Win XP |
| Windows XP Professional |
| Windows XP |
| windows 8 | Win 8 |
| Windows 8 |
| Windows 8 Pro |
| Windows 8.1 |
| Linux Ubuntu 10.04.4 LTS | Ubuntu |
| Ubuntu 13.04 |
| Windows Vista Home | Win Vista |
| Windows Vista |

Variable formato:

|  |  |
| --- | --- |
| doc | .doc |
| doc 97-2003 |
| Doc |
| docx | .docx |
| Docx |
| HTM | .html |
| html |
| HTML |
| abw | .abw |
| PDF | .pdf |
| pdf |
| odt | .odt |
| RTF | .rtf |
| (.RTF) |
| .rft |
| rtf |
| Documento xml | .xml |