# Análisis y Visualización de un Archivo Shape.

José Maria Castellanos Hernández (1), Andrea Joselín Martínez Sánchez(2).

(1) Universidad de Colima, Laguna de Juluapan #410 Fracc. Carlos de la Madrid, [jcastellanos@ucol.mx](mailto:jcastellanos@ucol.mx) . (2) Universidad de Colima, Acacia #1315 Fraccionamiento La Reserva, 28984, amartinez39@ucol.mx

Resumen

Se pretende desarrollar un programa que cumpla con los aspectos solicitados por el maestro con el fin de fortalecer la forma de que el alumno pueda integrar en un solo proyecto todos sus conocimientos, para ello se creará un código en lenguaje Python 2.7.

La idea de llevar a cabo este proyecto surgió de la incertidumbre de tener un archivo SHP sin poder analizarlo de una manera más eficiente toda la información contenida en el mismo y el poder visualizarlo directamente una gráfica con un ploteo de los elementos. Esperamos obtener resultados favorables desde que el código que generemos funcione a la perfección, así como que generemos lo establecido con anterioridad como objetivos.

**Palabras clave**: Shapefile, Python 2.7, Código.

Abstract

It is intended to develop a program that meets the aspects requested by the teacher in order to strengthen the student's way of being able to integrate all their knowledge into a single project, for which a code in Python 2.7 language will be created.

The idea of ​​carrying out this project arose from the uncertainty of having a SHP file without being able to analyse it in a more efficient way all the information contained in it and the power to visualize it directly a graph with a plot of the elements. We hope to obtain favourable results from the moment that the code we generate works perfectly, as well as that we generate what has been previously established as objectives.

**Keywords**: Shapefile, Python 2.7, Code

## INTRODUCCIÓN.

Dentro del programa de estudio de la materia programación ll está contemplado el desarrollo de un proyecto final en el cual se pongan en práctica todos aquellos conocimientos adquiridos en la materia de programación l y ll.

En las clases de programación ll se ha trabajado y analizado archivos SHP mediante códigos en el lenguaje Python, es por ello que nuestro equipo determino que se realizara un proyecto el cual permita a cualquier persona pueda examinar por medio de un análisis estadístico los datos contenidos en un archivo Shape y también se tiene contemplado generación de un ploteo que le dé la posibilidad al usuario visualizar por medio de la graficación de la información contenida en el SHP y así complementar dicho análisis.

Es importante mencionar que con el archivo SHP contendrá información de vías de comunicación del Estado de Colima, el cual se descargará de la página oficial del INEGI, dentro del mismo contiene datos geoespaciales y datos tabulares, estos últimos son los datos que se le aplicará el análisis, con la finalidad de proporcionar información valiosa tal como:

* Tipo de vía de comunicación.
* Cantidad de vectores existentes.
* Máximo.
* Mínimo.
* Medio.

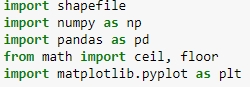
Gracias a este análisis se le dará a conocer al interesado toda la información relevante para la toma de decisiones importantes ya sea para la creación de más vías, mantenimiento o presupuestos.

## DESARROLLO.

Para poder lograr nuestros objetivos se tuvo que proceder a ejecutar una investigación con la finalidad de contar con las herramientas necesarias para concretar el proyecto.

Gracias a nuestra investigación determinamos que ocuparemos seis librerías las cuales nos permitirá realizar los objetivos del proyecto:

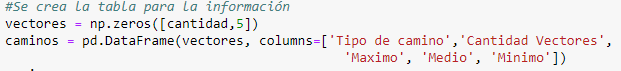
* Shapefile.
* Pandas.
* Numpy.
* Math.
* Matplotlib.pyplot



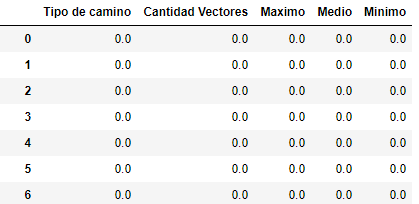
Cada una de ellas aplicara una proceso al archivo SHP, el primer objetivo es obtener la información del mismo para ello es necesario importar nuestro archivo para concretar esta acción entra en juego la librería “Shapefile” el cual nos otorga la facilidad de importar el archivo y a la vez leerlo con el comando “**sf = shapefile.Reader("Archivo.shp")**”



Para poder representar los datos que se quieren entregar es necesario la creación de una tabla. Dicha tabla contendrá tanto los datos de **entrada** como los de **salida**, los cuales serían los resultados del análisis estadístico que se le aplicara al archivo Shape.



Con este código se creara la tabla con 5 columnas, con los números de filas que contenga nuestro archivo SHP.



El análisis estadístico constará de los siguientes puntos:

* Cantidad de Vectores existentes
* Máximo
* Mínimo
* Medio

Y para realizada la tabla se precederá a ejecutar el análisis estadístico se utilizara la librería Math la cual permitirá realizar el procesamiento de los datos de nuestro archivo.

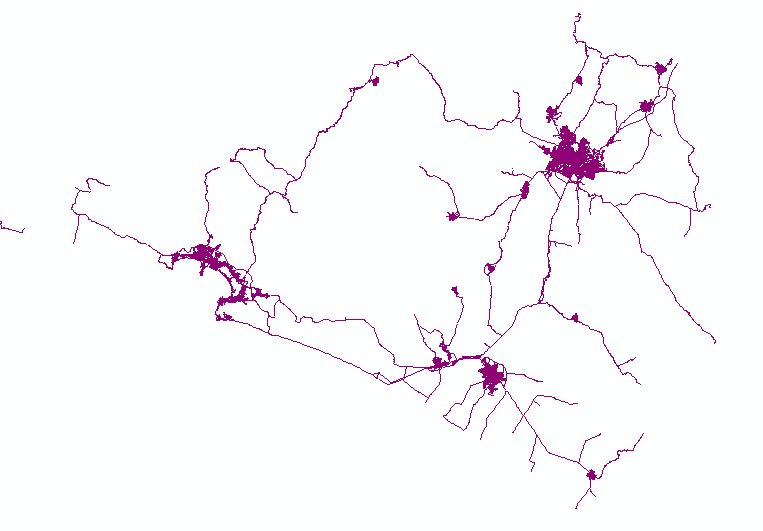
Se procederá al ordenamiento de los elementos esto con la finalidad de aplicar los ciclos que permitirá a analizar los puntos que se están evaluando.



El último objetivo que se quiere lograr es la generación del Ploteo el cual permitirá visualizar por medio de una gráfica la localización de cada uno de los elementos existentes en el archivo SHP, para lograr este objetivo entra en juego la librería Matplotlib.pyplotl el cual nos permitirá la obtención de las coordenadas geográficas (longitud, latitud) y así poderlas graficarlas.

## MANEJO DE DATOS

Para poder desarrollar el proyecto es necesario la utilización de archivos de entrada para poder aplicarles un proceso para poder obtener los archivos de salida que en este caso serían nuestros productos. En este caso práctico el archivo de entrada que se utilizará será un archivo Shape que lleva por nombre “red\_de\_carreteras\_col” el cual contiene todas las vías carreteras del estado de colima.



*Fig. Archivo Shape de las carreteras del estado de Colima.*

En este caso el tipo de vector encontrado en el mismo es de tipo de línea ya que son vialidades. Este archivo se obtuvo de la página oficial del “INEGI”.

Como ya sabemos los archivos Shape aparte de contar con información geoespacial cuentan con información tabular, en dicha información contiene distintos datos como los que se aprecian a continuación:

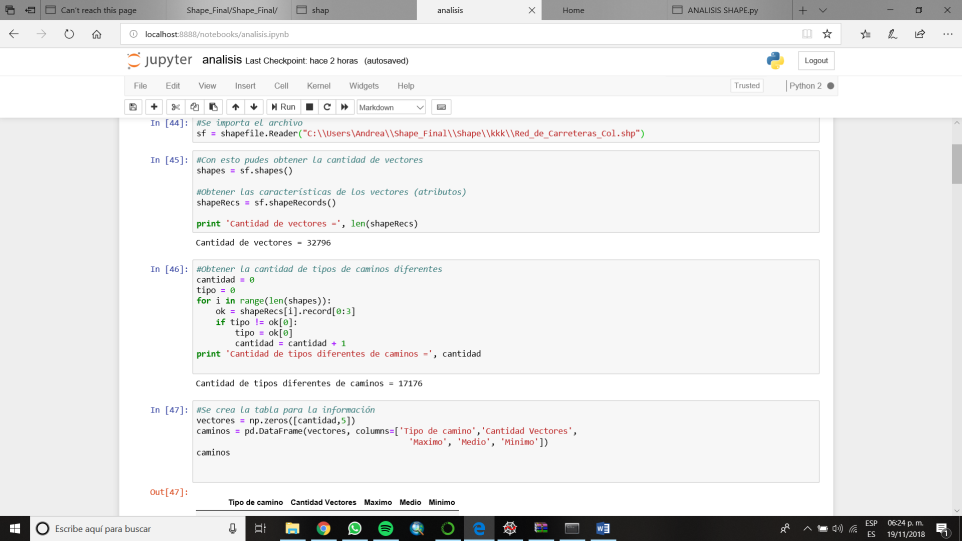


*Fig. Tabla de atributo del archivo Shape.*

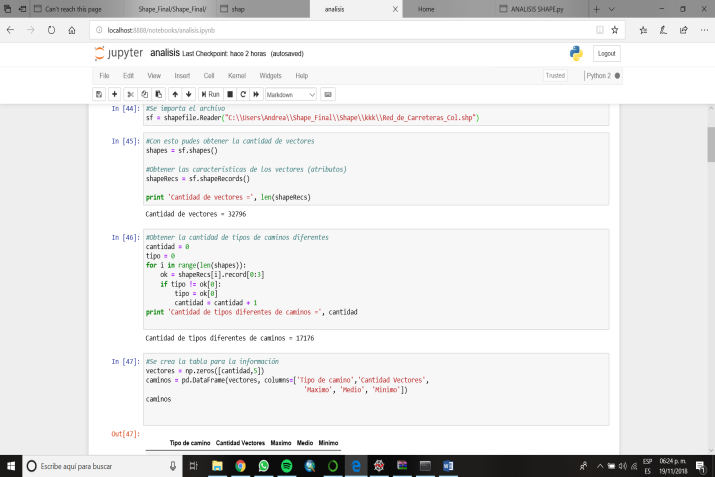
Contiene el tipo de vector que es, también nombre y tipo de vialidad y lo más relevante para este proyecto la longitud de cada una de las vialidades. Gracias a este tipo de información se puede hacer un procesamiento masivo de datos ya que como se puede observar es de todo el estado de Colima, la finalidad de este proceso es arrojar de una maneja rápida y eficiente el desglosamiento de toda ese conjunto de datos y así poder otorgarle al usuario las herramientas necesarias para la toma de decisiones.

### **RESULTADOS.**

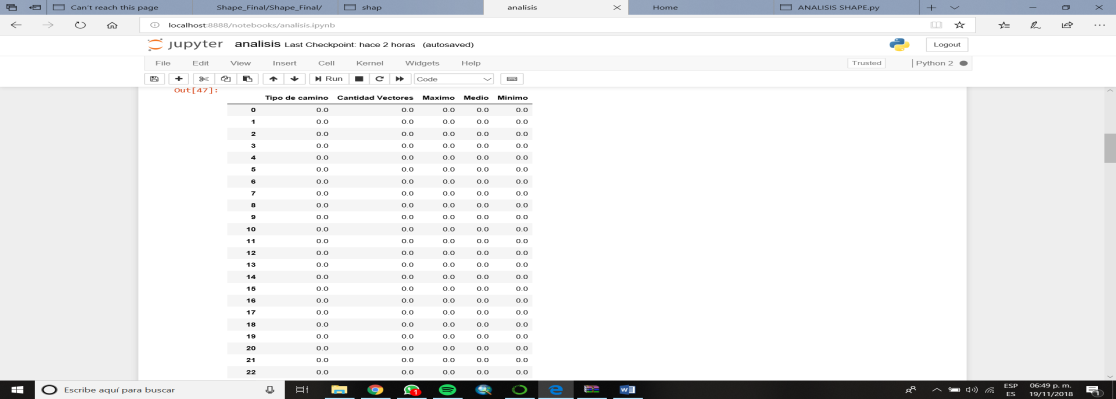
En base a los datos que se manejaron para llevar a cabo la realización del proyecto se obtuvo un código el cual cumpliera con los objetivos planteados con anterioridad con el fin de tener resultados favorables. Gracias a la buena elección de las librerías empleadas se pudo llevar a cabo, sin embargo. Los resultados de importar un archivo SHP fueron las siguientes:



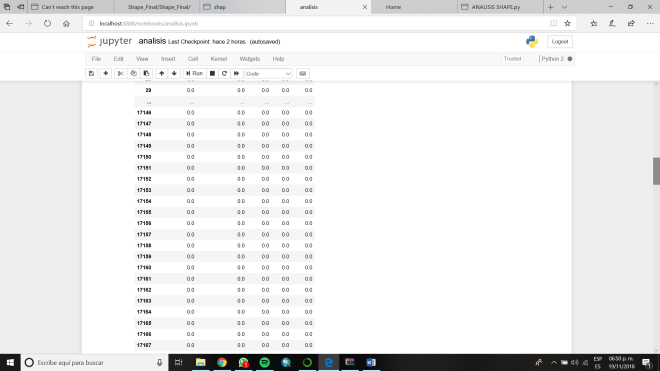
*Fig. Se muestra la cantidad de vectores con los que cuenta el archivo Shape.*



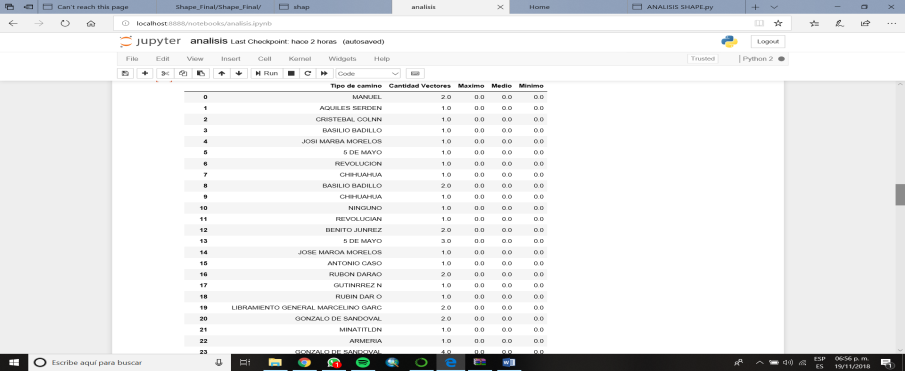
*Fig. Muestra los tipos diferentes de caminos que tiene el archivo Shape.*



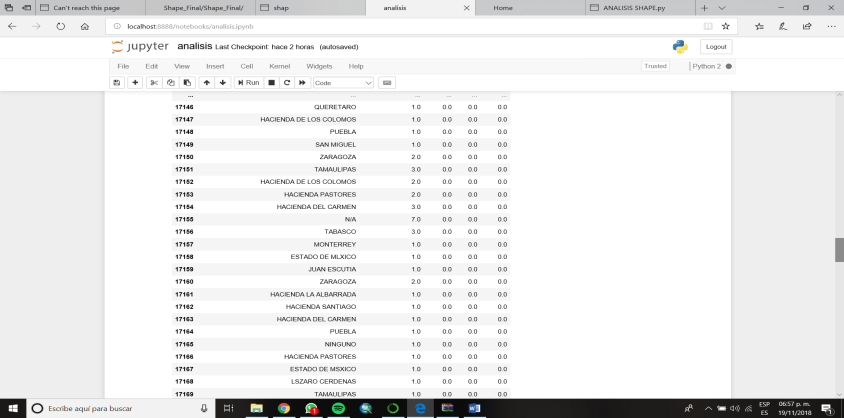
*Fig. Tabla para la información del archivo Shape.*



*Fig. Continuación de la tabla.*

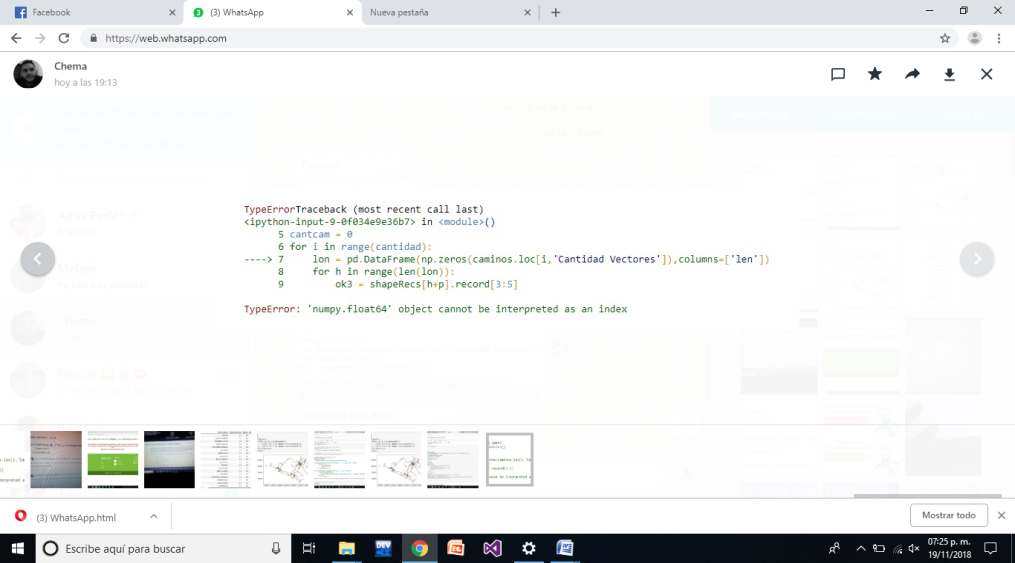


*Fig. Tabla con el tipo de camino.*



*Fig. Continuación de la tabla de tipo de caminos.*

Enseguida de la tabla con los tipos de camino con los que cuenta el archivo shape en el apartado para ordenar las longitudes se presentó el siguiente error:

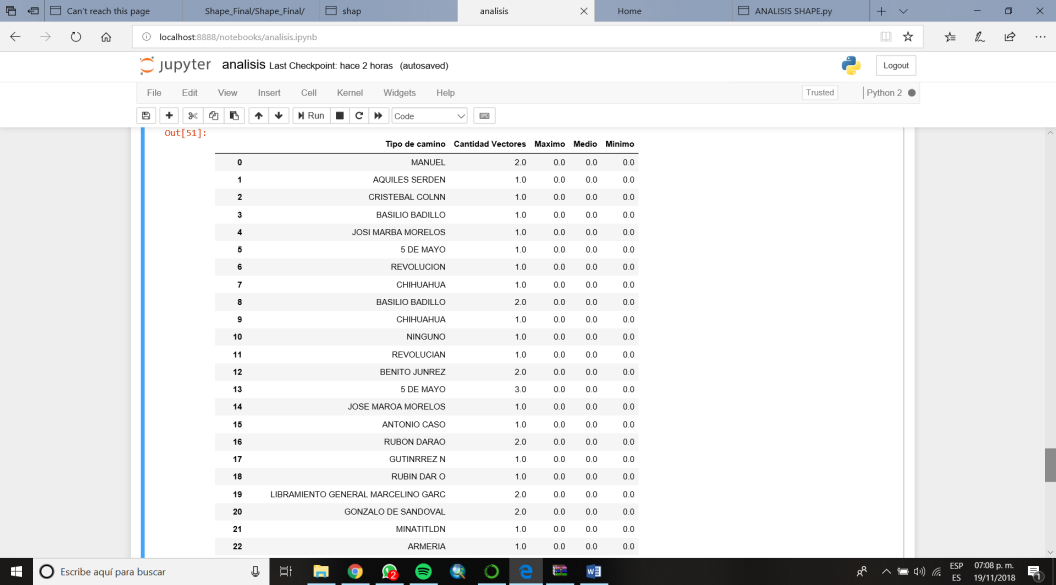


*Fig. Error generado en el código.*

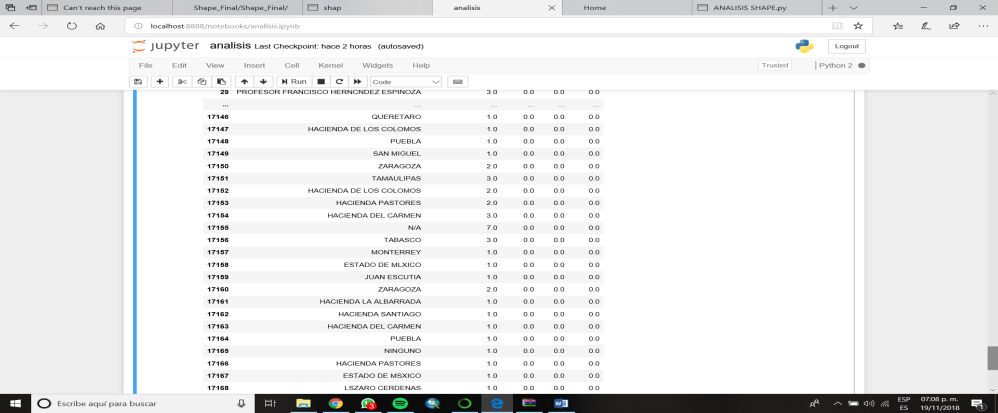
Ante este error, se investigó el motivo que lo ocasionaba; la investigación arrojo que era porque se requería una versión más actualizada de la paquetería de numpy. Con el siguiente comando se pretendía que se actualizará y corriera el código:

sudo pip install -U numpy==1.11.0

Sin embargo, no funciono.



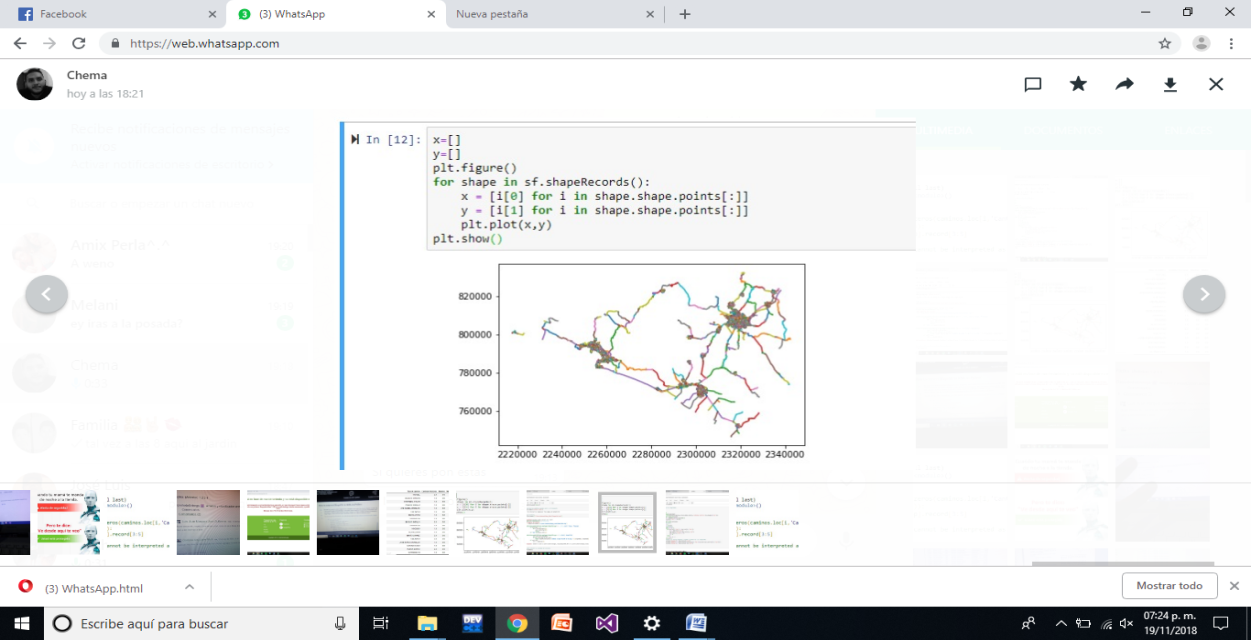
*Fig. Tabla donde se debe de presentar el analisis estadístico.*



*Fig. Continuación de la tabla.*

Debido al error planteado con anterioridad se puede observar que si afecto ya que no se pudo obtener el proceso estadístico del archivo Shape con el que trabajamos es por lo que se vuelve a mostrar en la columna de máximos, mínimos y medio como cero. Se logró alcanzar la mayoría de lo que se tenía planeado para el proyecto, faltaría indagar más para poder solucionar el problema del proceso estadístico.

Después del problema presentado se optó por desarrollar el ploteo del archivo Shapefile, cual arrojo el ploteo de los datos del archivo Shape para poderlo visualizar.



*Fig. Ploteo de los datos.*

### **CONCLUSIONES.**

Basándonos en los resultados que obtuvimos cuanto al código así como al manejo de datos podemos concluir que se cumplieron en su mayoría los objetivos aunque el código en un apartado nos marcó error sin poder generar el análisis estadístico que era en parte de gran importancia para el proyecto, se indago en fuentes de internet sin embargo no logramos que se resolviera esa situación, como se mencionó con anterioridad el posible problema se encontraba en la librería Numpy que su versión era desfavorable para lo que pretendíamos realizar. Respecto a lo demás del código podemos decir que si cumplió nuestras expectativas, dándonos el número de vectores, tipos de caminos, ploteo y demás que planteamos en él.

En general, el desarrollo del proyecto de desarrollo respecto a lo marcado en el programa, se tuvieron dificultades las cuales se trataron de dar una solución, es importante recalcar que se aprendió mucho con el manejo de datos de gran volumen y la visualización de los mismos, nos llevamos más conocimientos que al comienzo del proyecto.

Figura 1. Fachada principal de la biblioteca

Si las figuras son más anchas de lo que permite el ancho de la columna, sitúelas al final del documento en la sección sin columnas. Especifique en el texto (Figura n, en apéndice).

### Tablas

Deben realizarse con el mismo procesador de textos que el artículo. En texto portarán epígrafes indicativos de su contenido precedidos de la palabra "Tabla" y un numeral romano correlativo (ejemplo, "por Tabla I. ").

*Fig. Error presentado en el código.*

Si las tablas son más anchas de lo que permite la columna, sitúelas al final del documento en la sección sin columnas. Especifique en el texto (Tabla n, en apéndice).

Utilice para la cabecera de la tabla el estilo “Tabla-Cabecera) y para el texto de las celdas normales el estilo “Tabla-Texto”. Para la leyenda de pie de tabla, use el estilo “Leyenda de figura o tabla”.

|  |  |
| --- | --- |
| Estilo | Utilización |
| Subrayados | No usar |
| Negritas | No usar |
| Versales | No usar |
| Cursivas | Términos en otras lenguas diferentes de la del texto  Títulos de obras  Énfasis |
| Mayúsculas | Solo acrónimos |

Tabla I. Estilos de carácter permitidos y su uso

### Notas

Las notas explicativas deben usarse excepcionalmente. No utilice el sistema automático de su procesador de textos. Refiéralas en texto tecleando su número secuencial entre paréntesis, e inclúyalas al final del trabajo en un apartado titulado "Notas" situado inmediatamente antes del dedicado a las "Referencias", en orden secuencial y precedida cada una de ellas de su correspondiente número entre paréntesis seguido de tabulador. Las citas bibliográficas se realizarán como en el resto del texto; y la referencia completa se incluirá en el apartado de "Referencias". Las páginas web se consideran referencias bibliográficas y deben tratarse como tales, evitando su cita como nota.

Aplique a los párrafos de notas el estilo “Notas”.

### Citas bibliográficas en texto

Las referencias bibliográficas en texto se denotarán por el apellido del autor, el año del trabajo y, si es necesario, el número de página. Si conviene, el nombre y el apellido del autor podrán quedar fuera del paréntesis. Cuando se sucedan seguidamente varias referencias al mismo autor y obra, se podrá usar la abreviatura ibidem. Si dos autores de la lista de referencias poseen el mismo primer apellido, se referirán por sus dos apellidos y, si tienen ambos apellidos iguales, por el nombre también. Si dos o más trabajos del mismo autor son del mismo año, se distinguirán mediante letras del alfabeto secuenciales en minúsculas pospuestas a la fecha sin espacio. Si hay varias referencias dentro del paréntesis, se separarán por punto y coma si son de diferentes autores, y coma si son del mismo.

Así, por ejemplo, "Es una teoría sustentada por varios autores (Gallego, 1975; Fernández y Alonso, 1993) […]". "Según Jaime Pérez (1993, p. 24) se establecen […]". "El mismo autor (ibidem, p. 27) hace constar la conveniencia de […]". "Se han detectado diferencias en la replicación de la experiencia (Menéndez, 1994a, 1994b; Menéndez y Alonso, 1997)".

Si incluye texto literal de algún autor, inclúyalo entre comillas.

Si el texto literal que está reproduciendo supera las dos o tres líneas, siga un procedimiento diferente: hágalo sin comillas en un párrafo aparte con el estilo “Cita”. Si lo muestra incompleto, incluya tres puntos entre corchetes para indicar la elipsis […].

### Formato de las referencias bibliográficas

Las referencias bibliográficas se incluirán por orden alfabético al final del artículo bajo el título "Referencias", aplicándoles el estilo “Referencias”. Deberán seguir la norma UNE 50104:1996/ISO 690:1987, y, en cuanto a la puntuación, aproximarse lo más posible a las ISBD abreviadas, aunque sin espacio delante de punto, coma y punto y coma. Las referencias a páginas web deben incluirse también en este apartado, no en el texto ni en nota. Refiera adecuadamente en el texto del articulo toda la bibliografía utilizada; no incluya en el apartado referencias documentos que no haya incluido en el texto. Se presenta a continuación el formato para los principales tipos de materiales, con los elementos que pueden darse o no en cursivas, y seguidos de un ejemplo.

### Monografías, congresos, informes, normas y tesis

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. Edición. Lugar de publicación: editor; lugar de publicación: editor, fecha. ISBN número. Tesis doctoral.

Delclaux, Isidoro; Seoane, Julio (1982). Psicología cognitiva y procesamiento de la información: teoría, investigación y aplicaciones. Madrid: Ediciones Pirámide, 1982.

#### Artículos de publicaciones periódicas

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. // Título de la publicación periódica. ISSN numero. Volumen:número (Fecha de publicación) primera página-última página.

Ellis, David (1992a). The physical and cognitive paradigms in Information Retrieval Research. // Journal of Documentation. 48:1 (March 1992) 45-46.

#### Capítulos de una monografía, de un informe o contribuciones a un congreso

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. // Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. Edición. Lugar de publicación: editor; lugar de publicación: editor, fecha. ISBN número. Primera página-última página.

Markey, Karen (1990). Keyword searching in an online catalog enhanced with a library classification. // Bengtson, Betty G.; Hill, Janet Swan (eds.). Classification of library materials: current and future potential for providing access. New York: Neal-Shuman Publishers, 1990. 99-125.

Smith, Ph. J.; Beghtol, C.; Fidel, R.; Kwasnik, B. H. (eds.) (1993). Proceedings of the 4th ASIS SIG/CR Classification Research Workshop: Columbus, OH, Oct.24, 1993. Silver Spring, MD.: American Society for Information Science, 1993.

#### Páginas web y recursos en línea

Se tratarán como monografías, artículos o capítulos según corresponda y se añadirá al final detrás de punto el URL y la fecha de consulta entre paréntesis:

Referencia. URL (Fecha de consulta).

Sagredo Fernández, Félix; Espinosa Temiño, María Blanca (2000). Del libro, al libro electrónico-digital. // Cuadernos de Documentación Multimedia. 9 (2000). http://www.ucm.es/info/multidoc/multidoc/revista/num9/cine/ sagredo.htm (2001-01-22).

## El resumen

Un resumen proporciona información sobre objetivos, alcance, metodología, resultados, conclusiones y recomendaciones. Por alcance se entiende la determinación de hasta qué punto se han tratado los objetivos. Los resultados, conclusiones y recomendaciones se distinguen unos de otros así: los resultados son las informaciones o datos específicos y concretos que se obtienen de la aplicación de la metodología; las conclusiones se derivan de los resultados bien al interpretarlos en el marco de conocimientos teóricos previos o bien al crear un marco de interpretación nuevo a partir de ellos a través de un proceso de abstracción; finalmente, las recomendaciones se refieren a los nuevos problemas y vías de estudio que plantean a la vista de los resultados y conclusiones.

El resumen se debe redactar de forma clara y concisa, sin redundancias ni perífrasis ni información difícil de interpretar, como neologismos o abreviaturas que no se desarrollan. Debe tener un mínimo de cien palabras y un máximo de doscientas cincuenta.

Es muy conveniente que su primera frase contenga en primer lugar la mención del tipo de documento. Por ejemplo, "Estado de la cuestión sobre la aplicación del modelo europeo de calidad a las bibliotecas universitarias", "Proyecto de aplicación del MARC21 a la catalogación de páginas web institucionales", etc.

## La redacción del artículo científico: algunos aspectos obvios frecuentemente descuidados

* *Sistematicidad*. El artículo científico debe redactarse de forma muy organizada, con una estructura textual firme. En los informes de resultados de proyectos, se suele utilizar el esquema objetivos-metodología-resultados- conclusiones- recomendaciones. En los estados de la cuestión se aborda el asunto en orden cronológico o bien sistemático, según la ontología propia del área científica en cuestión; sigue la discusión, la síntesis, las conclusiones y las recomendaciones. Los artículos críticos suelen seguir el esquema tesis-síntesis-antítesis. Otros tipos documentales pueden exigir otras estructuras textuales, pero, en cualquier caso, dicha estructura debe ser explícita y aparente.
* *Claridad y concisión*. El lenguaje del artículo científico debe ser claro y conciso. Excepcionalmente, en los trabajos de tipo ensayístico las personas con un alto dominio del lenguaje escrito pueden permitirse la utilización de recursos literarios.
* *Pertinencia y parsimonia*. Debe evitarse proporcionar información ya conocida, salvo si constituye el punto central de la discusión. A ella cabe referirse por medio de las citas o, si estas no son necesarias, suponerla patrimonio de la comunidad profesional y científica. Sistematizar conocimiento es una función que se realiza solo en un tipo de artículo científico -la revisión o estado de la cuestión-; en los demás casos es contraproducente, aunque se deben citar las fuentes y las obras clave en la introducción del trabajo.
* *Civilidad*. Es necesario citar a otros autores que han trabajado sobre el tema, demostrando que se conoce la literatura previa y que el trabajo se inserta civilizadamente en un determinado contexto científico.

## Notas

(1) A propuesta de los autores o del comité científico, una versión ampliada de los trabajos presentados podrá ser remitida para su publicación en la revista *Scire: representación y organización del conocimiento* o en la revista Organización del Conocimiento en Sistemas de Información y Documentación tras recibir las sugerencias del comité científico, si las hubiere. Los trabajos no será publicados si no cuentan con el visto bueno de los evaluadores seleccionados en virtud de su especialización temática, que supervisarán los artículos basándose en el sistema de doble ciego.

## Referencias

Delclaux, Isidoro; Seoane, Julio (1982). Psicología cognitiva y procesamiento de la información: teoría, investigación y aplicaciones. Madrid: Ediciones Pirámide, 1982.

Ellis, David (1992a). The physical and cognitive paradigms in Information Retrieval Re search. // Journal of Documentation. 48:1 (March 1992) 45-46.

Markey, Karen (1990). Keyword searching in an online catalog enhanced with a library classification. // Bengtson, Betty G.; Hill, Janet Swan (eds.). Classification of library materials: current and future potential for providing access. New York: Neal-Shuman Publishers, 1990. 99-125.

Sagredo Fernández, Félix; Espinosa Temiño, María Blanca (2000). Del libro, al libro electrónico-digital. // Cuadernos de Documentación Multimedia. 9 (2000).

Smith, Ph. J.; Beghtol, C.; Fidel, R.; Kwasnik, B. H. (eds.) (1993). Proceedings of the 4th ASIS SIG/CR Classification Research Workshop: Columbus, OH, Oct.24, 1993. Silver Spring, MD.: American Society for Information Science, 1993.

## Apéndice

Utilice esta página y las siguientes para situar tablas y figuras de tamaño superior al ancho de la columna, u otros materiales. Si no la utiliza, borre esta sección.