**UNIVERSIDAD NACIONAL JOSÉ MARÍA ARGUEDAS**

**FACULTAD DE INGENIERÍA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

|  |
| --- |
| **TOPICOS ESPECIALES** |

**DOCENTE** : Ing. Ivan Soria Solis

**ESTUDIANTE** : Mayuri Ccahuana Huamaní

**ANDAHUAYLAS – APURÍMAC**

**PERÚ**

**OCTUBRE, 2015**

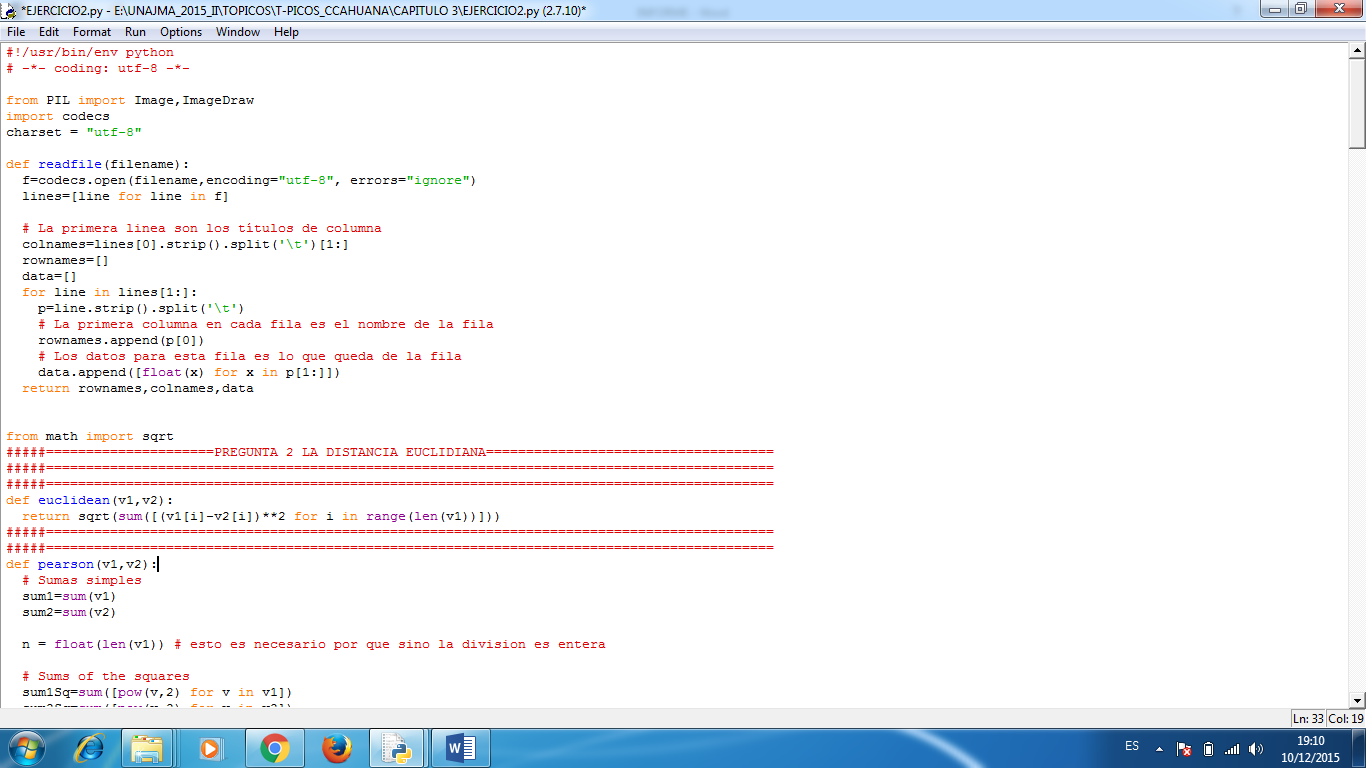
**PREGUNTA2:**

**Pruebe usando la distancia euclidiana para clustering de los blogs ¿Cómo cambia esto los resultados?**

En [matemáticas](https://es.wikipedia.org/wiki/Matem%C3%A1ticas), la distancia euclidiana o euclídea es la [distancia](https://es.wikipedia.org/wiki/Distancia) "ordinaria" (que se mediría con una regla) entre dos puntos de un [espacio euclídeo](https://es.wikipedia.org/wiki/Espacio_eucl%C3%ADdeo), la cual se deduce a partir del [teorema de Pitágoras](https://es.wikipedia.org/wiki/Teorema_de_Pit%C3%A1goras).

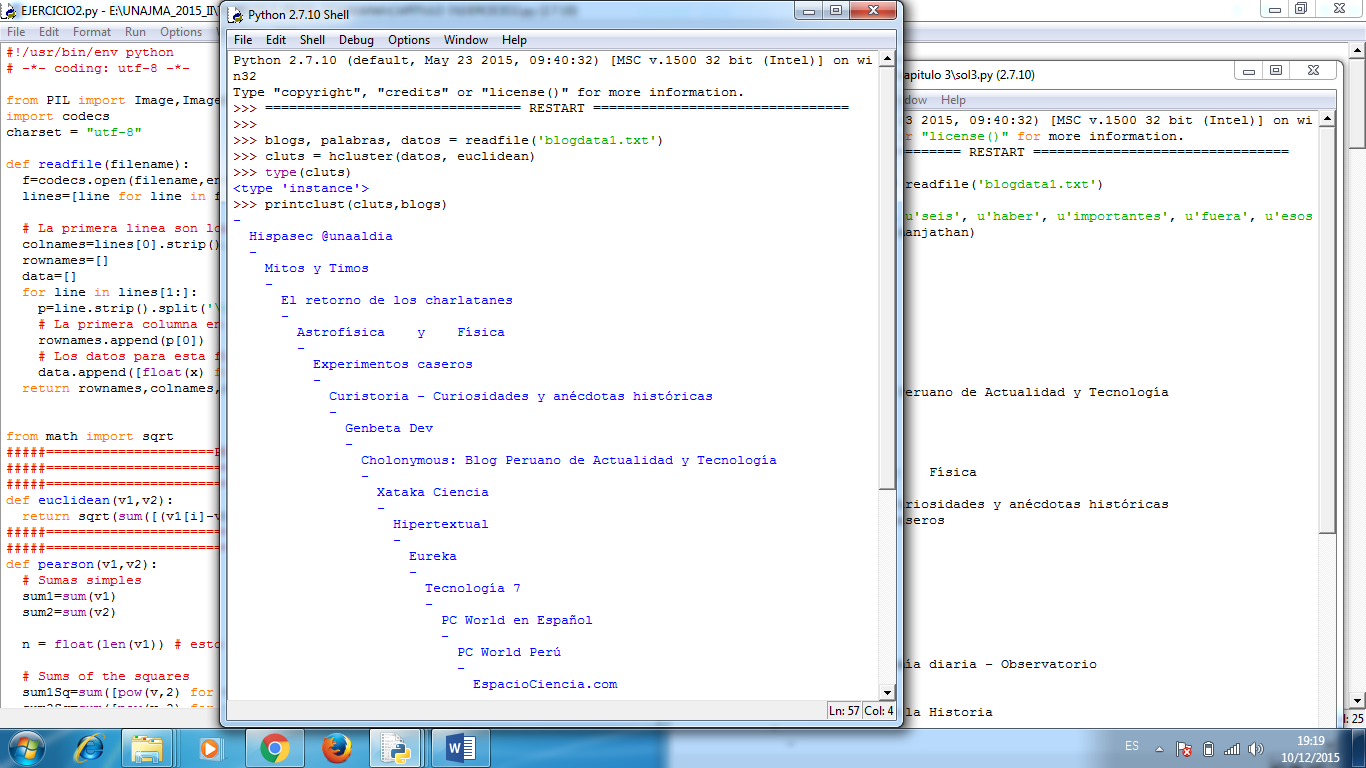
Por ejemplo, en un espacio bidimensional, la distancia euclidiana entre dos puntos *P*1 y *P*2, de [coordenadas cartesianas](https://es.wikipedia.org/wiki/Coordenadas_cartesianas) (*x*1, *y*1) y (*x*2, *y*2) respectivamente, es:

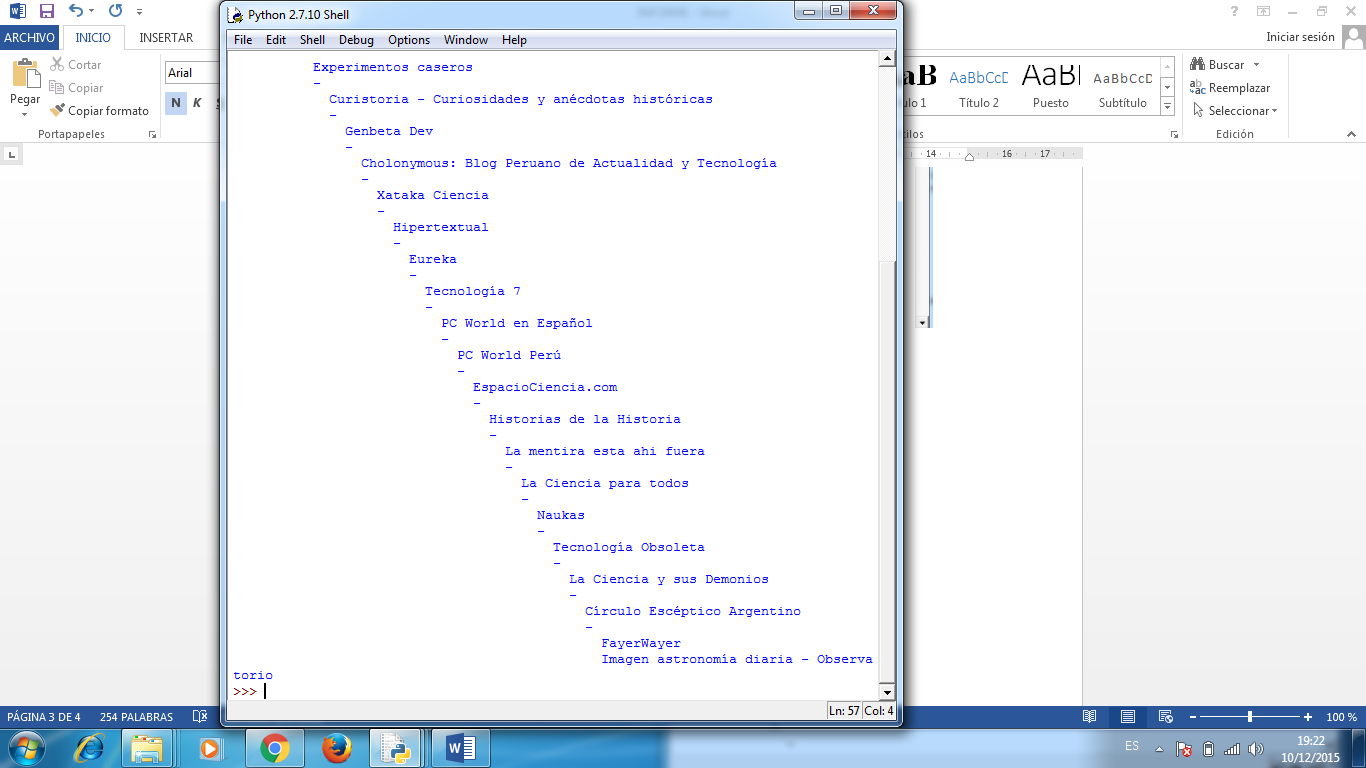
d_E(P_1,P_2)=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}

**Teniendo la siguiente formula nos vamos al código de clustering.**

AQUÍ SE TIENE LA FORMULA EUCLIDIANA

PARA VER LOS RESULTADOS VAMOS A EJECUTAR LA FUNCION (euclidean)

**Después de ejecutar nos muestra los siguientes resultados:**



**PREGUNTA2:**

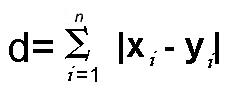
**Investigue a cerca de la distancia de Manhattan. Cree una función para esta y vean cómo cambian los resultados?**

**MANHATTAN**

**Descripción general**

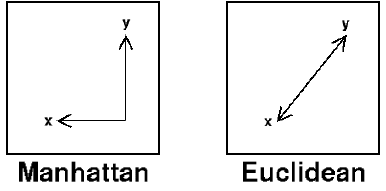
La función de la distancia Manhattan calcula la distancia que puede recorrer para llegar de un punto de datos a la otra si un camino en forma de rejilla es seguido. La distancia Manhattan entre dos elementos es la suma de las diferencias de sus correspondientes componentes.

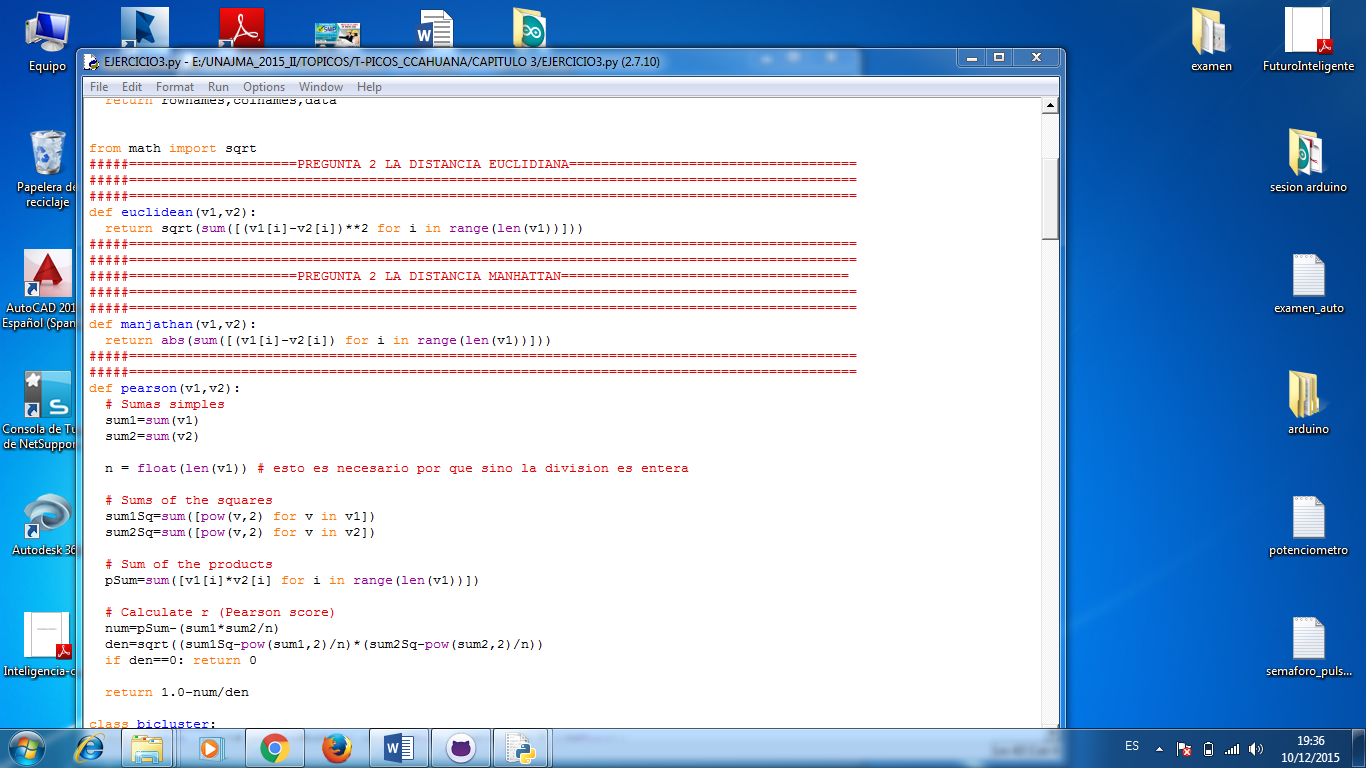
La fórmula para esta distancia entre un punto ***X*** = ***(X***1, ***X*** *2, etc.) y un punto****Y*** = ***(Y***1, ***Y*** *2, etc.) es:*



Donde n es el número de variables, y ***X****i* y ***Y****i* son los valores de la *i* ª variables, en los puntos ***X*** e ***Y, respectivamente.***

La siguiente figura ilustra la diferencia entre Manhattan distancia y distancia euclídea:



 T**eniendo la siguiente formula nos vamos al código de clustering.**

AQUÍ SE TIENE LA FORMULA DE LA DISTANCIA DE MANHATTAN

PARA VER LOS RESULTADOS VAMOS A EJECUTAR LA FUNCION (manjathan)

**Después de ejecutar nos muestra los siguientes resultados:**