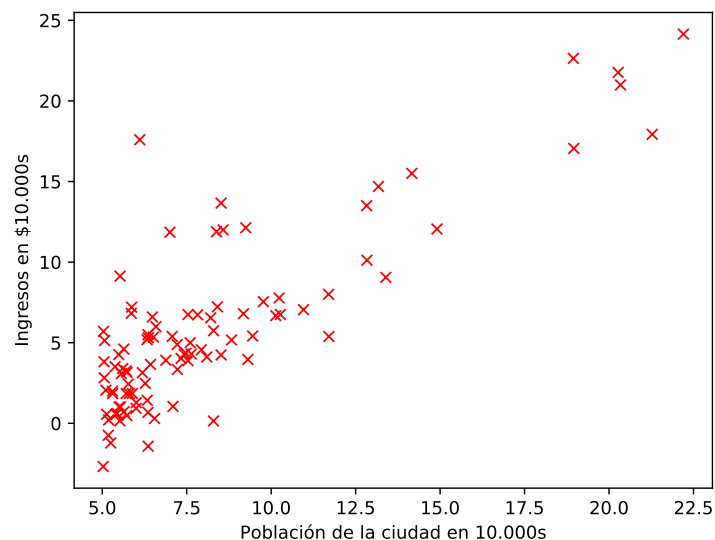

Práctica 1. Regresión lineal

Material proporcionado:

Fichero	Explicación
ex1data1.csv	Datos sobre los que aplicar regresión lineal con una variable.

1. Regresión lineal con una variable

En esta práctica has de aplicar el método de regresión lineal sobre los datos del fichero `ex1data1.csv` que representan datos sobre los beneficios (segunda columna en el archivo) de una compañía de distribución de comida en distintas ciudades, en base a su población (primera columna en el archivo), como se muestra en esta figura:



Para leer el contenido del archivo csv puedes utilizar el método `read_csv` de Pandas, que devuelve un *DataFrame*, y luego convertirlo a un array de numpy con el método `to_numpy()`:

```
1 import numpy as np
  from pandas.io.parsers import read_csv
3
valores = read_csv("ex1data1.csv", header=None).to_numpy()
```

Para encontrar los parámetros que definen la recta que mejor se ajusta a los datos de entrenamiento, has de aplicar el método de descenso de gradiente. El objetivo de la regresión lineal es minimizar la función de coste $J(w, b)$

En el método de descenso de gradiente nos vamos acercando iterativamente al valor de que minimiza la función de coste $J(w, b)$ actualizando cada componente.

A modo de depuración es bueno que vayas visualizando por pantalla el valor de la función de coste J a medida que avanza el descenso de gradiente, para comprobar que efectivamente su valor decrece de manera continua. Para ello has de implementar una función auxiliar que calcule el coste. El coste inicial con $w = 0$ y $b = 0$ es aproximadamente de 32.07.

Aplicando el método de descenso de gradiente, con unas 1500 iteraciones y un valor de $\alpha = 0.01$ deberías obtener unos parámetros tales que, por ejemplo, para una población de 70.000 habitantes ($x = 7$) predican unos beneficios aproximados de \$45.282 ($y = 4.5282$), y que definen una recta como la que se muestra en esta figura:

