# 15 Generar datos



*Data visualization* es el uso de representaciones visuales para explorar y presentar patrones en conjuntos de datos. Está estrechamente relacionado con *data analysis*, que utiliza código para explorar los patrones y conexiones de un conjunto de datos. Un conjunto de datos puede ser una pequeña lista de números que cabe en una sola línea de código, o pueden ser terabytes de datos que incluyen muchos tipos diferentes de información.

Crear visualizaciones de datos eficaces es algo más que hacer que la información tenga un aspecto agradable. Cuando la representación de un conjunto de datos es sencilla y visualmente atractiva, su significado resulta claro para los espectadores. La gente verá patrones y significados en tus conjuntos de datos que no sabían que existían.

Afortunadamente, no necesitas un superordenador para visualizar datos complejos. Python es tan eficiente que, con un simple ordenador portátil, puedes explorar rápidamente conjuntos de datos que contengan millones de puntos de datos individuales. Estos puntos de datos no tienen por qué ser números; con las nociones básicas que aprendiste en la primera parte de este libro, también puedes analizar datos no numéricos.

La gente utiliza Python para trabajos intensivos en datos en genética, investigación climática, análisis político y económico, y mucho más. Los científicos de datos han escrito una impresionante serie de herramientas de visualización y análisis en Python, muchas de las cuales también están a tu disposición. Una de las herramientas más populares es Matplotlib, una biblioteca de trazado matemático. En este capítulo, utilizaremos Matplotlib para hacer trazados sencillos, como gráficos de líneas y de dispersión. Luego crearemos un conjunto de datos más interesante basado en el concepto de paseo aleatorio: una visualización generada a partir de una serie de decisiones aleatorias.

También utilizaremos un paquete llamado Plotly, que crea visualizaciones que funcionan bien en dispositivos digitales, para analizar los resultados de tirar dados. Plotly genera visualizaciones que cambian automáticamente de tamaño para adaptarse a diversos dispositivos de visualización. Estas visualizaciones también pueden incluir una serie de características interactivas, como resaltar aspectos concretos del conjunto de datos cuando los usuarios pasan el ratón por encima de distintas partes de la visualización. Aprender a utilizar Matplotlib y Plotly te ayudará a empezar a visualizar los tipos de datos que más te interesan.

[Subtema 1 de 35: (Ver todo)](c15.html)[siguiente](c15_2.html)