

Instituto Politécnico Nacional
Centro de Investigación en Computo

Python en el Ámbito Científico

Profesor: Alan Badillo Salas

Sesión 3

— Problemas Científicos con Python

Hoja de Actividades y Ejercicios

Julio 5, 2025.

Actividad 5

El histograma es una herramienta útil para medir la frecuencia de los datos que caen sobre los intervalos que dividen el espacio de un eje de datos. Se construye dividiendo el eje de datos entre el valor mínimo observado y el valor máximo, en exactamente k regiones (intervalos o *bins*). Un *bin* es un intervalo descrito por el valor inferior y hasta antes de llegar al siguiente valor superior que parte el espacio. En cada región o intervalo se cuentan los datos que caen en ese intervalo para la serie de datos.

Crea una serie de datos con al menos 20 observaciones para una característica de interés sobre la muestra (de la población). Describe quién es la población, quién es la muestra y quién es la serie de datos. Luego determina el histograma, es decir, cuántos datos hay en cada *bin* o región, usa 4 o 5 regiones para dividir el espacio.

- Define una serie de datos con al menos 20 elementos, por ejemplo, edades.
- Describe de qué muestra viene esa serie de datos, por ejemplo, personas en un grupo de baile.
- Describe qué representa la población, por ejemplo, todas las personas que están aprendiendo a bailar.
- Ubica el valor mínimo y el máximo, por ejemplo, mínimo 18 y máximo 94.
- Divide el espacio en 4 intervalos mutuamente excluyentes y consecutivos, por ejemplo, de 18 a 36, de 37 a 55, de 56 a 74 y de 75 a 94.

Pista: usa la división del máximo menos el mínimo entre el número de intervalos $\frac{94-18}{4} = 19$ y genera los puntos de corte 18, 37, 56, 75 y 94.

- Cuenta el número de elementos que caen en cada intervalo, incluso podría ser cero si ningún dato de la serie cae en ese intervalo.
- Grafica el eje de datos dividido en los intervalos contra el número de puntos en cada intervalo, usa una barra para rellenar el intervalo.
- ¿Qué puedes observar en la gráfica de histograma?
- ¿En qué intervalos se concentra la mayor cantidad de datos?
- ¿Cómo podrías usar esta información para vender un producto o tomar alguna decisión respecto a la serie de datos?

Actividad 6

Responde las siguientes preguntas acerca de los problemas vistos en clase, indica la opción más adecuada:

1. ¿Qué es el Cuartil 1 y cómo se obtiene?
 - (a) Es el valor que representa el dato ordenado al 25% de la población y se obtiene ordenando los datos y seleccionando el del índice más próximo al 25% de la progresión.
 - (b) Es la cuarta parte entre el valor máximo y mínimo y se obtiene calculando el 25% del máximo menos el mínimo.
 - (c) Es el valor que toma la suma acumulada para una proporción del 25%, esta se calcula calculando las sumas acumuladas y midiendo la proporción, cuando la proporción llega al 25% el cuartil 1 representa la suma acumulada hasta ese punto.
2. ¿Qué es el Cuartil 2 y cómo se obtiene?
 - (a) Es la cuarta parte entre el valor máximo y mínimo y se obtiene calculando el 50% del máximo menos el mínimo.
 - (b) Es el valor que toma la suma acumulada para una proporción del 50%, esta se calcula calculando las sumas acumuladas y midiendo la proporción, cuando la proporción llega al 50% el cuartil 1 representa la suma acumulada hasta ese punto.
 - (c) Es el valor que representa el dato ordenado al 50% de la población y se obtiene ordenando los datos y seleccionando el del índice más próximo al 50% de la progresión.
3. ¿Qué es el Cuartil 3 y cómo se obtiene?
 - (a) Es el valor que toma la suma acumulada para una proporción del 75%, esta se calcula calculando las sumas acumuladas y midiendo la proporción, cuando la proporción llega al 75% el cuartil 3 representa la suma acumulada hasta ese punto.
 - (b) Es el valor que representa el dato ordenado al 75% de la población y se obtiene ordenando los datos y seleccionando el del índice más próximo al 75% de la progresión.
 - (c) Es la cuarta parte entre el valor máximo y mínimo y se obtiene calculando el 75% del máximo menos el mínimo.

Hoja 6 de Ejercicios

Crea un programa que defina dos series de datos: 10 estaturas y 10 pesos. Luego grafica un punto tipo estrella entre cada estatura (coordenada X) y cada peso (coordenada Y) y omite la línea que una los puntos.

```
estaturas = [  
    1.72, 1.68, 1.56, 1.84, 1.55,  
    1.67, 1.61, 1.73, 1.89, 1.51]  
  
pesos = [74, 67, 58, 95, 49,  
    65, 70, 81, 83, 52]  
  
import matplotlib.pyplot as pyplot  
  
pyplot.plot(estaturas, pesos,  
    marker="*", linestyle="none")
```

¿Crees que los datos tengan un comportamiento lineal o cuadrático? ¿Cómo lo justificarías?

Hoja 7 de Ejercicios

Crea un programa que defina una lista x con los números del 1 al 10 y una lista vacía y . Para cada x_i en x calcula $y_i = x_i^2$ y agrega y_i al final de la lista y . Luego grafica ambas coordenadas.

```
import matplotlib.pyplot as pyplot  
  
x = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]  
y = []  
  
for xi in x:  
    yi = xi ** 2  
    y.append(yi)  
  
pyplot.plot(x, y, marker=".", linestyle="none")
```

¿La gráfica es cuadrada o está mal distribuida? ¿Es correcto comparar el espacio x del 0 al 10 y el espacio y del 0 al 100?

Hoja 8 de Ejercicios

Crea un programa que defina una lista x con los números en el rango del 0 al 39 (tamaño 40) y una lista vacía y . Para cada x_i en x calcula $y_i = \sin(\frac{x_i}{20}\pi)$ y agrega y_i al final de la lista y .

```
import matplotlib.pyplot as pyplot
from math import sin, pi
```

```
x = list(range(40))
y = []
```

```
for xi in x:
    yi = sin(xi / 20 * pi)
    y.append(yi)
```

```
pyplot.plot(x, y, marker=".", linestyle="--")
```

¿Qué representa un *linestyle* de --? ¿Qué pasa si se divide entre 40 en lugar de 20? ¿Qué pasa si se divide entre 10 en lugar de 20?

Hoja 9 de Ejercicios

Crea un programa que defina una lista t a partir de un rango de tamaño 100 (de 0 a 99) y dos listas vacías x y y . Para cada t_i en t , calcula $x_i = \cos(\frac{t_i}{50}\pi)$ y $y_i = \sin(\frac{t_i}{50}\pi)$, dibuja las coordenadas x y y cambiando el color de la línea y el color del marcador.

```
import matplotlib.pyplot as pyplot
from math import cos, sin, pi
```

```
t = list(range(100))
```

```
x = []
y = []
```

```
for ti in t:
    xi = cos(ti / 50 * pi)
    x.append(xi)
    yi = sin(ti / 50 * pi)
    y.append(yi)
```

```
pyplot.plot(x, y, linestyle="-", color="grey",
            marker=".", markerfacecolor="red", markersize=10)
```

¿Qué figura se dibuja? ¿Usan solo 40 puntos y la división por 20 qué observas?

Hoja 10 de Ejercicios

Crea un programa que defina una lista t a partir de un rango de tamaño 1,000 (de 0 a 999) y dos listas vacías x y y . Para cada t_i en t , calcula $x_i = t_i \cdot \cos(\frac{t_i}{100}\pi)$ y $y_i = t_i + \sin(\frac{t_i}{100}\pi)$, dibuja las coordenadas x y y en color gris punteado.

```

import matplotlib.pyplot as pyplot
from math import cos, sin, pi

t = list(range(1_000))

x = []
y = []

for ti in t:
    xi = ti * cos(ti / 100 * pi)
    x.append(xi)
    yi = ti + sin(ti / 100 * pi)
    y.append(yi)

pyplot.plot(x, y, linestyle=":", color="grey")

¿Qué figura se forma? ¿Qué pasa si inviertes el * y el + entre  $x_i$  y  $y_i$ ?

```

Reto de la semana

Haz un programa que solicite tres coordenadas de los puntos $A = (a_x, a_y)$, $B = (b_x, b_y)$ y $C = (c_x, c_y)$ y dibuja el triángulo descrito.

Indica los vértices del triángulo de color rojo y el borde del triángulo de color azul.