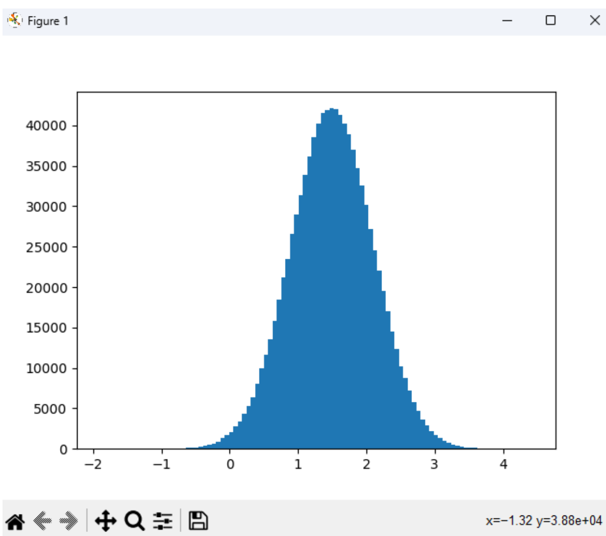
Jared Kinneer
4396828

1. Predict travel times
    a. Distance, time of day, day of the week, weather, destination zip code, start zip code
    b. Time (in hours and minutes)
    c. You could collect data from mapping apps (Waze, Google Maps, Apple Maps), or you could ask people to self-report.
    d. You would need a massive amount of data to make a connection between all of the different variables, and the correlation between different variables may be different based on location
2. At risk for hair loss
    a. Age, gender, family history of baldness
    b. At risk for hair loss (Yes/No)
    c. Collect at doctor's office, or places that consultations for hair loss treatment
    d. Different things cause hair loss in different people, might be hard to track all potential sources

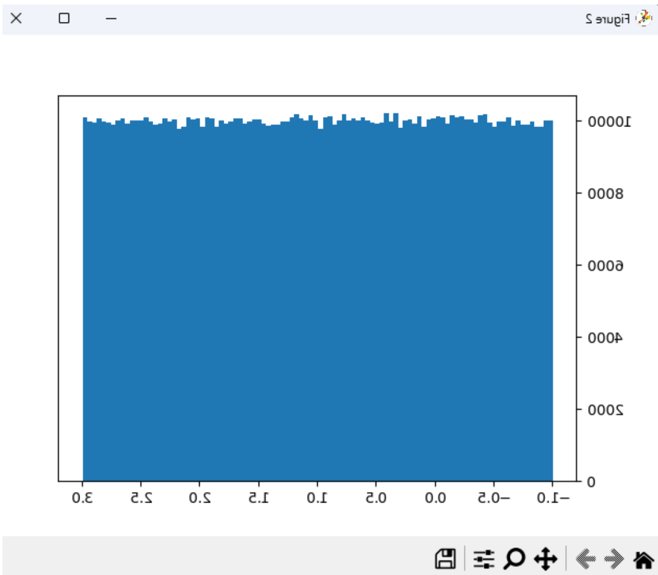
3c: The distribution of x does look Gaussian, as it takes the shape of a bell. And z does look normal, as it takes the shape of a rectangle

ps1-3-c-1



Jared Kinneer
4396828

ps1-3-c-2



3d: elapsed time: 1.5740067958831787 seconds

3e: elapsed time vector addition: 0.0019388198852539062 seconds, and vector addition is more efficient for adding a constant

3f: first run count= 624560, second run count= 624936

There is a difference because z is randomly generated so the exact amount of numbers in its range less than 1.5 will differ each time it is generated

4c: x=3/10 y=4/10 z =0

to normalize a vector sum the square of all components, then take the square root

norm $x1 = \sqrt{.5^2 + 0^2 + (-1.5)^2} = 1.58$

norm $x2 = \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2} \approx 1.41$

numpy results

$||x1|| = 1.5811$

$||x2|| = 1.4142$

Jared Kinneer
4396828

5:

First test array [[1,2,3], [4,5,6]], result [17, 29, 45]

Second test array [[1,2,3], [4,5,6] , [7,8,9]] result [66, 93, 126]