Exercise 1: Pandas and Numpy #Importing the required libraries import numpy as np import matplotlib.pyplot as plt import pandas as pd import matplotlib.pyplot as plt **Matrix Multiplication** To create a numpy matrix X of dimensions $n \times m$, where n = 100 and m = 20. Initialize Matrix X with random values. Then to create a numpy vector Y of dimensions m \times 1. Initialize the vector Y with values from a normal distribution using $\mu = 2$ and $\sigma = 0.01$ X=np.random.randint(10,size =(100,20)) # initializing X with random values #Initializing vector Y with the mean and std given Y=np.random.normal(2,0.01,(20,1)) print("Y vector-",Y) Y vector- [[2.01000581] [2.00726404] [1.98657501] [2.01951683] [1.99636613] [1.98879839] [1.98743152] [1.99813393] [2.00009748] [2.00239709] [2.00766363] [1.98906483] [1.97472742] [1.97709336] [1.99021826] [1.98266885] [1.99235989] [2.0068848] [1.98881392] [2.00478144]] print("X matrix-",X) print("\nShape of X:", X.shape) X matrix- [[7 1 8 ... 0 1 9] [2 4 9 ... 4 2 8] [9 6 2 ... 9 7 0] [7 1 2 ... 3 9 1] [2 4 5 ... 7 7 0] [1 8 2 ... 7 6 1]] Shape of X: (100, 20) To iteratively multiply (element-wise) each row of the matrix X with vector Y and sum the result of each iteration in another vector C. In [4]: C= np.zeros(100) #Initialize vector C for i in range(100): #for each row in matrix sum=0 for j in range(20): #for each column in the row sum += X[i][j]* Y[j] # multiplying the required elements and adding them to the previous value to get sC[i] = sum# storing the sum of the element wise multiplication of a row print("Value of Vector C: ", C) print("\nShape of vector C: ",C.shape) Value of Vector C: [169.53826628 179.47048188 179.87307917 197.79865232 193.3427728 169.59957395 143.65645206 191.57583815 197.55025403 163.61980952 193.89712075 183.48844036 157.62613832 187.65068285 235.5222964 225.32873432 129.72596375 197.38872024 167.68677748 163.33038359 159.58569034 161.61832615 205.78520145 219.70539728 231.54898159 163.52403256 181.43324894 151.49153916 189.71152758 187.50560027 137.65753738 163.30770068 211.54388739 177.31304166 201.76864024 171.10098336 171.41487113 141.63395496 203.90382745 195.51118964 169.57177795 191.48877752 201.56327364 193.69237433 177.69313089 187.43336237 199.61501659 161.56966324 175.40113267 201.56893834 187.47860964 147.67260725 197.57866356 187.44870576 227.60922086 151.62432979 177.68565965 177.4577411 163.60066097 159.37342826 191.31559879 177.66330284 177.49977323 141.59172136 189.30276042 161.84783869 161.44213589 221.63728464 139.64235813 167.59589802 207.5844346 181.57412204 179.50238655 187.57287835 205.511776 195.68177793 177.76114528 181.52297667 191.59175957 185.7443218 163.33796465 197.63073902 205.61437728 171.65100274 141.63215605 197.7230497 155.7030299 195.63512215 153.71877722 169.60863844 229.69191531 157.565221 177.6542577 165.58663796 117.70904857 173.86317469 197.58873911 161.61126047 151.59724585 143.7130126] Shape of vector C: (100,) To find mean and standard deviation of the new vector c. Here I am using the numpy functions for the same. #Finds average of all the values in the vector mean = np.mean(C)print("Mean of vector C:", mean) Mean of vector C: 179.44284313045046 # Finds the variation of the values in the vector from its mean value print("Standard Deviation of vector C:",std) Standard Deviation of vector C: 23.355834124633724 Plotting the histogram of vector C using 5 bins In [9]: plt.hist(C, edgecolor="red", bins=5) plt.title("Histogram of vector C") plt.xlabel("The range of values in vector") plt.ylabel("The number of values ") plt.show() Histogram of vector C 35 30 The number of values 25 20 15 10 5 220 240 120 140 160 180 200 The range of values in vector We observe that the histogram plot is similar to that of normal distribution. GRADING Performing tasks on the csv file given. import pandas as pd #Reading the required csv files using pandas functionality data frame = pd.read csv('Downloads/Grades.csv') data frame **First Name** Last Name English Maths Science German Sports Final Grade Hobgood 60.95 24.77 20.60 69.32 8.36 184.00 Robyn Swearngin 100.00 100.00 1 Eddy 100.00 12.99 52.24 365.23 2 83.37 100.00 78.69 100.00 19.50 Leoma Bridgman 381.56 3 87.75 100.00 86.93 87.90 41.73 404.31 Arnetta Peart 4 Maryland Colby 100.00 100.00 100.00 18.87 88.72 407.59 5 93.93 Sherron Sherron 92.06 55.91 -56.74 77.71 262.87 100.00 6 78.26 100.00 25.60 100.00 403.86 Glendora Christopher 7 100.00 64.53 100.00 23.21 79.01 366.75 Darlena Gunn 8 100.00 83.49 100.00 100.00 92.32 475.81 Aldo Armas 9 Tiny Jack 94.35 33.09 82.57 31.13 100.00 341.14 10 100.00 5.54 33.82 12.07 Carlton Elms 36.52 187.95 11 Herbert 50.73 -0.10 67.76 100.00 55.98 274.37 Lauretta Dimond 80.37 100.00 69.02 100.00 79.62 429.01 Almeta 100.00 13 Phoebe Schill 70.37 100.00 47.00 77.37 394.74 14 Paris 18.75 73.80 87.00 59.30 100.00 338.85 Krystyna 100.00 15 Laffoon 100.00 100.00 34.98 94.55 429.53 Miyoko 16 Rebecca 70.79 97.81 52.25 19.76 -13.93 226.68 Duck 17 Elwanda Hyland 45.69 74.86 43.10 45.00 76.72 285.37 Kerrick 68.70 75.87 -14.87 57.32 84.28 18 Gretchen 271.30 19 Winnifred Colonna 83.79 100.00 85.66 3.94 100.00 373.39 20 100.00 100.00 100.00 99.52 62.59 462.11 Gidget Casseus 100.00 80.29 100.00 21 Elaina Mcdougal 100.00 -36.81 343.48 22 100.00 100.00 100.00 100.00 455.10 Shoshana Goldberger 55.10 23 100.00 100.00 40.44 100.00 100.00 440.44 Argentina Nelson 24 Lyle Millsaps 100.00 71.88 100.00 -17.33 100.00 354.55 55.90 25 Janay Julius 41.70 100.00 75.53 100.00 373.13 18.61 26 Devorah Heyden 0.84 50.14 83.58 65.68 218.85 27 72.41 61.00 100.00 76.38 361.16 Thelma Romberger 51.37 28 35.40 66.92 69.67 71.08 -2.34240.73 Armanda Hendley 29 Raymon Myerson 49.83 27.36 61.90 72.97 13.11 225.17 Compute the sum for all subjects for each student. #Finding sum along the subject columns data frame['sum'] = data frame[['English','Science','German','Sports',' Maths']].sum(axis=1) data frame Last Name English Maths Science German Sports Final Grade First Name sum 0 69.32 Robyn Hobgood 60.95 24.77 20.60 8.36 184.00 184.00 1 Eddy 100.00 12.99 100.00 52.24 100.00 365.23 365.23 Swearngin 2 83.37 100.00 78.69 100.00 19.50 381.56 Leoma Bridgman 381.56 3 Arnetta 87.75 100.00 86.93 87.90 41.73 404.31 404.31 4 100.00 100.00 100.00 18.87 88.72 407.59 407.59 Maryland Colby 5 55.91 93.93 262.87 262.87 Sherron Sherron 92.06 -56.74 77.71 6 100.00 25.60 100.00 100.00 403.86 403.86 Glendora Christopher 78.26 7 Darlena Gunn 100.00 64.53 100.00 23.21 79.01 366.75 366.75 8 Aldo Armas 100.00 83.49 100.00 100.00 92.32 475.81 475.81 9 Tiny 94.35 33.09 82.57 31.13 100.00 341.14 341.14 Jack 10 100.00 36.52 5.54 33.82 12.07 187.95 187.95 Carlton Elms 11 67.76 100.00 55.98 274.37 Lauretta Herbert 50.73 -0.10 274.37 69.02 12 80.37 100.00 100.00 79.62 429.01 429.01 Almeta Dimond 13 Phoebe Schill 100.00 70.37 100.00 47.00 77.37 394.74 394.74 14 18.75 73.80 87.00 59.30 100.00 338.85 338.85 Krystyna Paris 100.00 100.00 100.00 429.53 429.53 15 Miyoko Laffoon 34.98 94.55 70.79 -13.93 226.68 226.68 16 Rebecca Duck 97.81 52.25 19.76 17 Elwanda Hyland 45.69 74.86 43.10 45.00 76.72 285.37 285.37 18 Gretchen Kerrick 68.70 75.87 -14.87 57.32 84.28 271.30 271.30 83.79 100.00 373.39 373.39 19 Winnifred Colonna 85.66 3.94 100.00 20 100.00 100.00 100.00 99.52 62.59 462.11 462.11 Gidget Casseus 21 Elaina Mcdougal 100.00 100.00 80.29 100.00 -36.81 343.48 343.48 22 Shoshana Goldberger 55.10 100.00 100.00 100.00 100.00 455.10 455.10 23 Argentina Nelson 100.00 100.00 40.44 100.00 100.00 440.44 440.44 24 100.00 71.88 100.00 -17.33 100.00 354.55 354.55 Lyle Millsaps 25 Julius 41.70 100.00 55.90 75.53 100.00 373.13 373.13 Janay 26 Devorah 0.84 18.61 50.14 83.58 65.68 218.85 218.85 Heyden 100.00 76.38 27 Thelma Romberger 72.41 61.00 51.37 361.16 361.16 28 Armanda Hendley 35.40 66.92 69.67 71.08 -2.34 240.73 240.73 29 Raymon 49.83 27.36 61.90 72.97 13.11 225.17 225.17 Myerson We observe that the column Final grade already contains the sum of all the subjects. data frame.columns Index(['First Name', 'Last Name', 'English', 'Maths', 'Science', 'German', 'Sports', 'Final Grade', 'sum'], dtype='object') • Compute the average of the point for each student(Total= 500) - 5 subjects are there In [14]: #Finding average of the subjects and assigning it to new column data frame['average'] = data frame[['sum']]/5 data frame Last Name English Maths Science German Sports Final Grade First Name sum average 0 Robyn 60.95 24.77 20.60 69.32 8.36 184.00 184.00 36.800 Hobgood 1 Eddy 100.00 12.99 100.00 52.24 100.00 365.23 365.23 73.046 Swearngin 2 83.37 100.00 78.69 100.00 19.50 381.56 381.56 76.312 Leoma Bridgman 3 86.93 404.31 Arnetta Peart 87.75 100.00 87.90 41.73 404.31 80.862 4 Colby 100.00 100.00 100.00 18.87 88.72 407.59 407.59 81.518 Maryland Sherron 262.87 5 Sherron 92.06 55.91 93.93 -56.74 77.71 262.87 52.574 100.00 403.86 403.86 6 25.60 100.00 100.00 80.772 Glendora Christopher 78.26 7 Darlena 64.53 100.00 366.75 366.75 Gunn 100.00 23.21 79.01 73.350 8 100.00 83.49 100.00 100.00 92.32 475.81 475.81 95.162 Aldo Armas 9 94.35 33.09 82.57 31.13 100.00 341.14 341.14 68.228 Tiny Jack 37.590 10 Carlton Elms 100.00 36.52 5.54 33.82 12.07 187.95 187.95 11 100.00 274.37 274.37 Lauretta Herbert 50.73 -0.10 67.76 55.98 54.874 80.37 100.00 69.02 12 Dimond 100.00 79.62 429.01 429.01 85.802 Almeta 13 Phoebe Schill 100.00 70.37 100.00 47.00 77.37 394.74 394.74 78.948 14 Krystyna Paris 18.75 73.80 87.00 59.30 100.00 338.85 338.85 67.770 100.00 100.00 15 Miyoko Laffoon 100.00 34.98 94.55 429.53 429.53 85.906 16 70.79 97.81 52.25 19.76 -13.93 226.68 226.68 45.336 Rebecca Duck 285.37 285.37 17 Elwanda Hyland 45.69 74.86 43.10 45.00 76.72 57.074 18 Gretchen Kerrick 68.70 75.87 -14.87 57.32 84.28 271.30 271.30 54.260 100.00 373.39 19 85.66 3.94 100.00 373.39 Winnifred Colonna 83.79 74.678 20 Gidget Casseus 100.00 100.00 100.00 99.52 62.59 462.11 462.11 92.422 21 100.00 100.00 80.29 100.00 -36.81 343.48 343.48 68.696 Elaina Mcdougal 55.10 100.00 100.00 22 Shoshana Goldberger 100.00 100.00 455.10 455.10 91.020 Argentina 100.00 100.00 40.44 100.00 100.00 440.44 440.44 23 Nelson 88.088 -17.33 354.55 354.55 24 100.00 71.88 100.00 100.00 70.910 Lyle Millsaps 373.13 373.13 25 Janay Julius 41.70 100.00 55.90 75.53 100.00 74.626 218.85 43.770 26 Devorah Heyden 0.84 18.61 50.14 83.58 65.68 218.85 27 100.00 Thelma Romberger 72.41 61.00 76.38 51.37 361.16 361.16 72.232 28 35.40 66.92 69.67 71.08 -2.34 240.73 240.73 48.146 Armanda Hendley 29 Raymon 49.83 27.36 61.90 72.97 13.11 225.17 225.17 45.034 Myerson • Compute the standard deviation of point for each student. #Here we find the variation of each student from the mean of the average marks of all students. Thus we can kno #each student is .ie, if they are above average or below average of the class. total avg = data frame['average'].mean() data frame['standard deviation'] = data frame['average'] - total avg data frame Last Name English Maths Science German Sports Final Grade sum average standard deviation First Name 0 Robyn Hobgood 60.95 24.77 20.60 69.32 8.36 184.00 184.00 36.800 -31.726867 100.00 365.23 365.23 1 Eddy Swearngin 100.00 12.99 52.24 100.00 73.046 4.519133 19.50 381.56 381.56 2 83.37 100.00 78.69 100.00 76.312 7.785133 Leoma Bridgman 3 Arnetta Peart 87.75 100.00 86.93 87.90 41.73 404.31 404.31 80.862 12.335133 Maryland 100.00 100.00 407.59 81.518 12.991133 4 Colby 100.00 18.87 88.72 407.59 262.87 5 Sherron Sherron 92.06 55.91 93.93 -56.74 77.71 262.87 52.574 -15.952867 78.26 100.00 6 25.60 100.00 100.00 403.86 403.86 80.772 12.245133 Glendora Christopher 7 366.75 366.75 Darlena 100.00 64.53 100.00 23.21 79.01 73.350 4.823133 Gunn 8 100.00 83.49 100.00 100.00 92.32 475.81 475.81 95.162 26.635133 Aldo Armas 94.35 33.09 82.57 31.13 100.00 341.14 Jack 341.14 68.228 -0.298867 100.00 187.95 36.52 5.54 33.82 10 Carlton Elms 12.07 187.95 37.590 -30.936867 11 Lauretta Herbert 50.73 -0.10 67.76 100.00 55.98 274.37 274.37 54.874 -13.652867 12 80.37 100.00 69.02 100.00 79.62 429.01 429.01 85.802 17.275133 Almeta Dimond Phoebe 100.00 394.74 394.74 78.948 13 Schill 100.00 70.37 47.00 77.37 10.421133 -0.756867 73.80 100.00 338.85 338.85 14 Krystyna 18.75 87.00 59.30 67.770 Paris 15 Miyoko Laffoon 100.00 100.00 100.00 34.98 94.55 429.53 429.53 85.906 17.379133 16 Rebecca Duck 70.79 97.81 52.25 19.76 -13.93 226.68 226.68 45.336 -23.190867 17 Elwanda Hyland 45.69 74.86 43.10 45.00 76.72 285.37 285.37 57.074 -11.452867 18 68.70 75.87 84.28 271.30 54.260 -14.266867 Gretchen Kerrick -14.8757.32 271.30 Colonna 19 Winnifred 83.79 100.00 85.66 3.94 100.00 373.39 373.39 74.678 6.151133 20 Gidget Casseus 100.00 100.00 100.00 99.52 62.59 462.11 462.11 92.422 23.895133 21 Elaina Mcdougal 100.00 100.00 80.29 100.00 -36.81 343.48 343.48 68.696 0.169133 22 100.00 100.00 100.00 100.00 455.10 22.493133 Shoshana Goldberger 55.10 455.10 91.020 23 100.00 100.00 40.44 100.00 100.00 440.44 440.44 88.088 19.561133 Argentina Nelson 24 100.00 71.88 100.00 -17.33 100.00 354.55 354.55 70.910 2.383133 Lyle Millsaps 25 Janay Julius 41.70 100.00 55.90 75.53 100.00 373.13 373.13 74.626 6.099133 218.85 26 18.61 50.14 83.58 65.68 218.85 43.770 -24.756867 Devorah Heyden 0.84 27 72.41 61.00 100.00 76.38 51.37 361.16 361.16 72.232 3.705133 Thelma Romberger 28 35.40 66.92 69.67 71.08 -2.34 240.73 240.73 48.146 -20.380867 Armanda Hendley 61.90 29 Raymon Myerson 49.83 27.36 72.97 13.11 225.17 225.17 45.034 -23.492867 • Plot the average points for all the students (in one figure). #Plotting the average marks of all the students fig = plt.figure() ax = fig.add axes([0,0,1,1])ind = np.arange(7)ax.set_ylabel('Average points of each student') ax.set xlabel('Last name of each student') ax.set title('Average points for all students') ax.plot(data frame["Last Name"], data frame["average"]) plt.xticks(rotation=90) plt.show() Average points for all students 90 Average points of each student 70 50 40 Hobgood
Sweamgin
Bridgman
Peart
Colby
Sherron
Christopher
Gunn
Armas Mcdougal -Goldberger -Nelson -Millsaps -Julius -Heyden -Romberger -Hendley Myerson Casseus Last name of each student For each student assign a grade based on the rubric given GRADING SYSTEM Grade # of students % Range - 100 0 86 89 80 85 1 79 0 69 0 0 59 56 0 55 0 0 In [19]: conditions = [(data frame['average'] < 56),</pre> (data_frame['average'] >= 56) & (data_frame['average'] < 60),</pre> (data_frame['average'] >= 60) & (data_frame['average'] < 66),</pre> (data_frame['average'] >= 66) & (data_frame['average'] < 70),</pre> (data_frame['average'] >= 70) & (data_frame['average'] < 76),</pre> (data_frame['average'] >= 76) & (data_frame['average'] < 80),</pre> (data_frame['average'] >= 80) & (data_frame['average'] < 86),</pre> (data_frame['average'] >= 86) & (data_frame['average'] < 90),</pre> (data_frame['average'] >= 90) & (data_frame['average'] < 96),</pre> (data_frame['average'] >= 96) & (data_frame['average'] <= 100)</pre> values = ['F', 'D', 'C', 'C+', 'B-', 'B', 'B+', 'A-', 'A', 'A+'] #grade values according to the order of the condit data frame['Grade'] = np.select(conditions, values) # assigning values to new column based on the conditions sg data frame Last Name English Maths Science German Sports Final Grade sum average standard deviation Grade First Name 0 Hobgood 60.95 184.00 Robyn 24.77 20.60 69.32 8.36 184.00 36.800 -31.726867 F 100.00 Eddy 100.00 12.99 100.00 52.24 365.23 365.23 73.046 4.519133 Swearngin 2 Bridgman 83.37 100.00 100.00 381.56 Leoma 78.69 19.50 381.56 76.312 7.785133 В 3 Peart 87.75 100.00 86.93 B+ Arnetta 87.90 41.73 404.31 404.31 80.862 12.335133 407.59 4 Maryland Colby 100.00 100.00 100.00 18.87 88.72 407.59 81.518 12.991133 Sherron Sherron 55.91 93.93 -56.74 262.87 -15.952867 100.00 100.00 6 Glendora Christopher 78.26 100.00 25.60 403.86 403.86 80.772 12.245133 B+ 7 100.00 64.53 100.00 79.01 366.75 366.75 73.350 4.823133 Darlena Gunn 23.21 B-8 Aldo 100.00 83.49 100.00 100.00 92.32 475.81 475.81 95.162 26.635133 Armas Α 9 94.35 33.09 82.57 31.13 100.00 341.14 341.14 68.228 -0.298867 Tiny Jack C+ -30.936867 10 Carlton Elms 100.00 36.52 5.54 33.82 12.07 187.95 187.95 37.590 F -13.652867 F 50.73 100.00 55.98 274.37 274.37 11 Lauretta Herbert -0.10 67.76 54.874 80.37 100.00 69.02 429.01 85.802 12 Almeta Dimond 100.00 79.62 429.01 17.275133 B+ Schill 100.00 70.37 100.00 47.00 77.37 394.74 394.74 78.948 10.421133 В 13 Phoebe 100.00 14 Krystyna Paris 18.75 73.80 87.00 59.30 338.85 338.85 67.770 -0.756867 C+ 17.379133 15 Miyoko Laffoon 100.00 100.00 100.00 34.98 94.55 429.53 429.53 85.906 B+ 16 Rebecca Duck 70.79 97.81 52.25 19.76 -13.93 226.68 226.68 45.336 -23.190867 F 17 45.69 74.86 43.10 45.00 76.72 285.37 285.37 57.074 -11.452867 D Elwanda Hyland Kerrick 84.28 271.30 271.30 -14.266867 F 18 Gretchen 68.70 75.87 -14.87 57.32 54.260 Winnifred 19 83.79 100.00 3.94 100.00 373.39 373.39 74.678 Colonna 85.66 6.151133 B-100.00 99.52 20 Gidget Casseus 100.00 100.00 62.59 462.11 462.11 92.422 23.895133 Α 100.00 100.00 100.00 21 Mcdougal 80.29 343.48 343.48 68.696 0.169133 Elaina -36.81 C+ 100.00 55.10 100.00 100.00 455.10 455.10 91.020 22 Shoshana Goldberger 100.00 22.493133 Α 23 100.00 100.00 100.00 100.00 440.44 440.44 88.088 19.561133 Argentina Nelson 40.44 Α-100.00 24 Millsaps 100.00 71.88 100.00 -17.33 354.55 354.55 70.910 2.383133 Lyle B-25 Janay Julius 41.70 100.00 55.90 75.53 100.00 373.13 373.13 74.626 6.099133 Devorah 65.68 218.85 218.85 43.770 -24.756867 F 26 Heyden 0.84 18.61 50.14 83.58 27 72.232 3.705133 Thelma Romberger 72.41 61.00 100.00 76.38 51.37 361.16 361.16 B-28 35.40 66.92 69.67 71.08 -2.34240.73 240.73 48.146 -20.380867 F Armanda Hendley 29 49.83 27.36 61.90 72.97 13.11 225.17 225.17 45.034 -23.492867 F Raymon Myerson Plot the histogram of the final grades. plt.hist(data frame["Grade"], edgecolor="red") plt.title("Histogram of the final grades") plt.xlabel("Grades") plt.ylabel("Number of students") plt.show() Histogram of the final grades Number of students 2 В B+ Grades References

https://datagy.io/pandas-conditional-column/ https://www.geeksforgeeks.org/bin-size-in-matplotlib-histogram/