Data Analysis in Python.

Anaconda

Anaconda Installed?

<https://www.anaconda.com/products/individual>

|  |  |
| --- | --- |
| **Name** | **Yes/No** |
| Reuel Mendoza | Yes |
| Chris Ellis | Yes |
| Ryker Kang | yes |
| Jason Goertz | Yes |
| Natalia Solar | yes |
| Devon Gronquist | no |
| Matt Smithers | Yes |
| Krissi Postor | yes |
| Han Ji | yes |
| JaeWook Jung | yes |
| Beni Ungur | Yes |
| Mark Ignatovich | Yes |
| Tamires B | Yes |
| Alex Inclan | Yes |
| Spencer H | Yes |
| William Judge | Yes |
| Bryan Aleman | Yes |
| Thomas Parcher | Yes |
| Alexk K | Yes |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Problem**

In Python get an array of all numbers from 1 to 10

Take that array and apply the function (2 \* x + 1) to each element

Sum all the elements in the array

[Marcelo] Paste the code instead of just the result

|  |  |
| --- | --- |
| Group | Code |
| 1 | array = [1,2,3,4,5,6,7,8,9,10]  sum = 0  for i in array:  sum += 2 \* i + 1  print(sum)  120 |
| 2 | numArray = list(range(1,11))  total = 0  for num in range(len(numArray)):  total += numArray[num] \* 2 + 1    print(total) |
| 3 | arr = [ i+1 for i in range(10) ]  arr = [ 2\*x+1 for x in arr ]  total = sum(arr)  print(total) |
| 4 |  |
| 5 | myarray = [1,2,3,4,5,6,7,8,9,10]  total = 0  for x in myarray:  x = (2\*x)+1  total += x  print(myarray)  print(total) |
| 6 | a\_list = list(range(1,11))  b\_list = [2 \* x + 1 for x in a\_list]  total = sum(b\_list)  print(total) |
| 7 | array = [1,2,3,4,5,6,7,8,9,10]  total = 0  for x in array:  total += (2 \* x + 1)  print(total) |

Part

With the help of Google

Solve first part of problem 1 in the Midterm

Read World Indicators File.

Filter by United States

Create a Linear Model of Population by Year

Predict Population in 2015

Paste your code on the table.

Your numbers should be exactly the same as the ones in the midterm

|  |  |
| --- | --- |
| Group | Code |
| 1 | import pandas  import numpy  from sklearn.linear\_model import LinearRegression  data = pandas.read\_csv('c:\\midterm\\World Indicators.csv')  dusa = data[data["Country/Region"] == "United States"]  y = list(dusa["Population Total"])  x = numpy.array(dusa["Year"]).reshape(-1, 1)  model = LinearRegression().fit(x, y)  print(model.predict(numpy.array([[2015]]))) |
| 2 | import numpy as np  import pandas as pd  from sklearn.linear\_model import LinearRegression  data = pd.read\_csv("WorldIndicators.csv")  newdata = data[data['Country/Region'] == "United States"]  model = np.polyfit(newdata["Year"], newdata['Population Total'], 1)  predict = np.poly1d(model)  year = 2015  predict(year) |
| 3 | import pandas  import numpy  from sklearn.linear\_model import LinearRegression  data = pandas.read\_csv('c:\\midterm\\World Indicators.csv')  dusa = data[data["Country/Region"] == "United States"]  y = list(dusa["Population Total"])  x = numpy.array(dusa["Year"]).reshape(-1, 1)  model = LinearRegression().fit(x, y)  print(model.predict(numpy.array([[2015]]))) |
| 4 |  |
| 5 | import pandas as pd  import os  import numpy as np  import statsmodels.api as sm  Pwd  variable = pd.read\_csv(r"C:\DataAnalytics\World Indicators.csv")  variable.head()  is\_usa = variable['Country/Region']=='United States'  print(is\_usa.head())  var\_usa = variable[is\_usa]  print(var\_usa)  x = list(var\_usa["Year"])  x = sm.add\_constant(x)  y = list(var\_usa["Population Total"])  model = sm.OLS(y,x).fit()  predictions = model.predict([1,2015])  model.summary()  print(predictions) |
| 6 |  |
| 7 |  |

**Back at 7:50PM**

**Problem**

Model for **Europe**

Life Expectancy Female vs GDP + Health Expense as %GDP, Infant mortality Life Expectancy Male

GDP = 1000000000,

Health.Exp...GDP = 0.04,

Infant.Mortality.Rate = 0.05,

Life.Expectancy.Male = 80

|  |  |
| --- | --- |
| Group | Code |
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 |  |
| 6 |  |
| 7 |  |