**Course:** [INSERT COURSE ID HERE]
**Instructor:** [INSERT INSTRUCTOR NAME HERE]
**Due Date:** [INSERT DUE DATE HERE]

**Homework #** [INSERT HOMEWORK NUMBER] **– Recidivism Score**

## 1. INTRODUCTION

Throughout the nation, judges, probation, and parole officers are growing increasingly reliant on algorithms that will determine a given defendant's likelihood of recidivism. Because of the growing popularity of these algorithms and the systemic issues that accompany them, you will be exploring the factors that are included to determine the likelihood of becoming a recidivist.

What is Recidivism?

Recidivism is defined as *the tendency of a convicted criminal to re-offend*. The criminal justice system uses recidivism algorithms to predict this tendency. An algorithm (a set of instructions that receive an input and provide an output to accomplish a task or solve a problem) is used to assess whether or not a person is likely to recidivate. This algorithm is used for deciding sentencing or bail depending on the state using this form of "recidivism profiling".

How it works

Factors about a person such as age, zip code, life situations, race, criminal history, and more are input into recidivism algorithms and scores are calculated as output. Some algorithms categorize output into designations such as "low risk", "medium risk", and "high risk".

How it fails

Recidivism algorithms, even when they do not collect information about a person's race, have been found to skew results out of favor for people of color more so than White people.

Problems presented

These risk assessment algorithms are used to determine sentencing and in some states (like California) are used in lieu of cash bail. If black people are considered higher risk more often than white people, the already problematic criminal justice system once again discriminates against people of color and puts them at a further disadvantage.

## 2. GETTING STARTED

Using the information above and your own research, you will build a very simplified version of a recidivism risk algorithm to calculate a score for different inmates. This will function as an introduction to topics in 'data types', 'conditionals', 'reading files,' and 'functions'.

Research has shown that these are some of the features that go into these algorithms. Your task is to construct a function that will take in seven parameters that represent these features. This function will calculate a recidivism risk score based on the parameter inputs.

Prospective recidivism data is organized in the **recidivism.csv** file such that the data for each student is on one line, with the values separated by tabs. The data is in the following order: [age, gender, education, employed, crimeHistory, zipCode, race].

An example of two offenders data might be:

*[21, M, SC, True, True, 80021, White or Caucasian]*

This would represent a 21-year-old man with some college education, who is employed and has some criminal history, lives in the 80021 zip code, and is White / Caucasian.

*[45, NB, CC, True, False, 80204, Asain or Pacific Islander]*

This would represent a 45-year-old who is non-binary, has a community college education, is employed and has no criminal history, lives in the 80204 zip code, and is Asain / Pacific.

Calculate_Score(age, gender, education, employed, crimeHistory, zipCode, race):

> *age:* An integer that will specify the person's age. This variable will have a minimum input of 15 and a maximum input of 70.
>
> *gender:* A string that specifies the gender of the person. This variable will either be {'M': Male, 'F': Female, 'NB': Non-Binary, 'O':Other}.
>
> *education:* A string that specifies the education level of the person. This variable will either be {'MSD': Middle School Degree, 'SHS': Some High School, 'HSD': High School Degree, 'SC': Some College, 'CC': Community College, 'UD': Undergraduate Degree, 'GD': Graduate Degree}.
>
> *employed:* A Boolean that is True if the person is employed and False if the person is not.
>
> *crimeHistory:* A boolean that is True if the person has a criminal history and False if the person is not.

*zipCode*: An integer that will represent zip codes in the Denver Metropolitan Area of each person.

*race:* A string that specifies the race of the person. This variable will be "White or Caucasian, "Hispanic/Latino", "Black or African American", "Asain or Pacific Islander", "Native American or Alaskan Native", "Multiracial or Biracial".

## 3. Your Task

You get to decide how the score is calculated under one assumption: a higher score means a person has a higher likelihood of committing a crime. This would imply that a higher score is more likely to create a greater sentence compared to a person with a low score. **You will have to justify the allocation of scoring for every parameter in the function. Please be sure to include documentation and outside resources through comments or Python markdown.** Should older inmates receive a higher score than younger inmates? That is up to you. **Please take your time when considering your scoring and outside research is heavily encouraged.** Creating helper functions is also encouraged but not required.

<u>Discussion Questions</u>

How can recidivism algorithms contain a racial bias when race is not collected?

Is it possible to create a recidivism algorithm without biases?

How will education affect an offender's score? How will employment affect an offender's score? Please explain why varying education backgrounds or employment history may lead to a higher or lower score.

What role does gender play in calculating a recidivism score?

What are examples of some factors that could decrease the likelihood of recidivism?

What are your thoughts on having a recidivism score algorithm integrated with our judicial system? Is there a case for having them in place? A case for removing them entirely? Is there a middle option? Explain your justification.