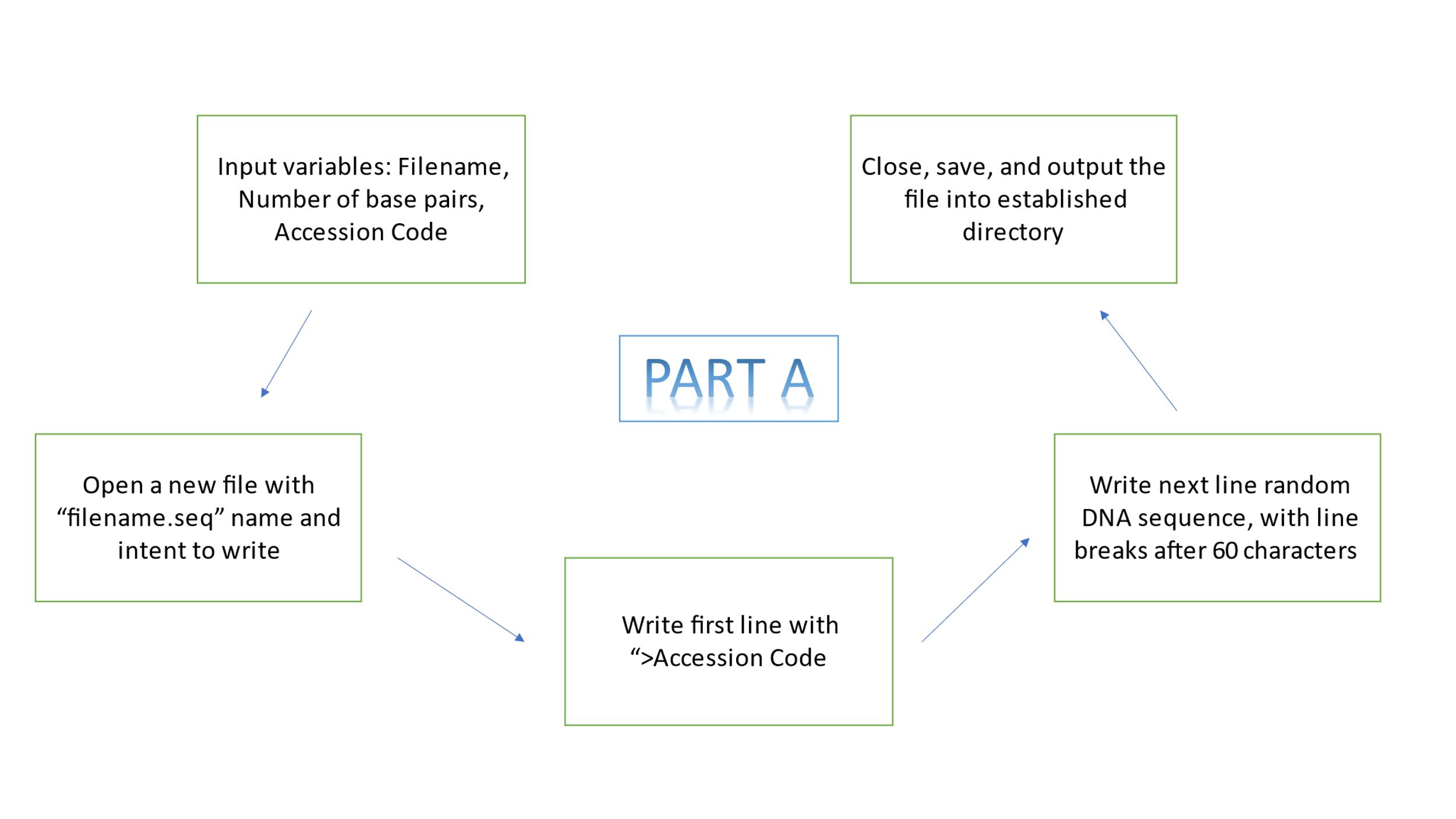
### Sanjana Venkatesh, sv2699 (github)

# Pipelines:

# Part A- The code I've written does not entirely work. I am not too comfortable with Python, and I couldn't figure out how to do it before the deadline.

![](.Data/pipeline\_images/Slide1.JPG):

# I don’t know why the above isn’t working in R markdown, I set the directory and everything!



# Part B- The code does work, I sampled it with a .txt file, I couldn’t try a .seq file, but I think the logic I’ve used makes sense, so it should work.

Set working directory, start

function with input

variable: Filename

Import Filename.seq

, and

convert the DNA sequence

alone to a data frame, for

easier use

Select each base and check

for repetitions, output will

be a list

Check for the length of the

list, for the count of each

base, and print the results

Output results to the

working directory with

appropriate Filename



PART B

# Part C- Because I couldn’t figure out Part A, I made a csv file with 4 columns and 100 random numbers in each to use for this. I have attached the csv file too for context.

