

# Homework 4 - BIOS 7649

Dominic Adducci

```
## Warning: package 'ggplot2' was built under R version 4.3.2

## Warning: package 'tidyr' was built under R version 4.3.2

## Warning: package 'readr' was built under R version 4.3.2

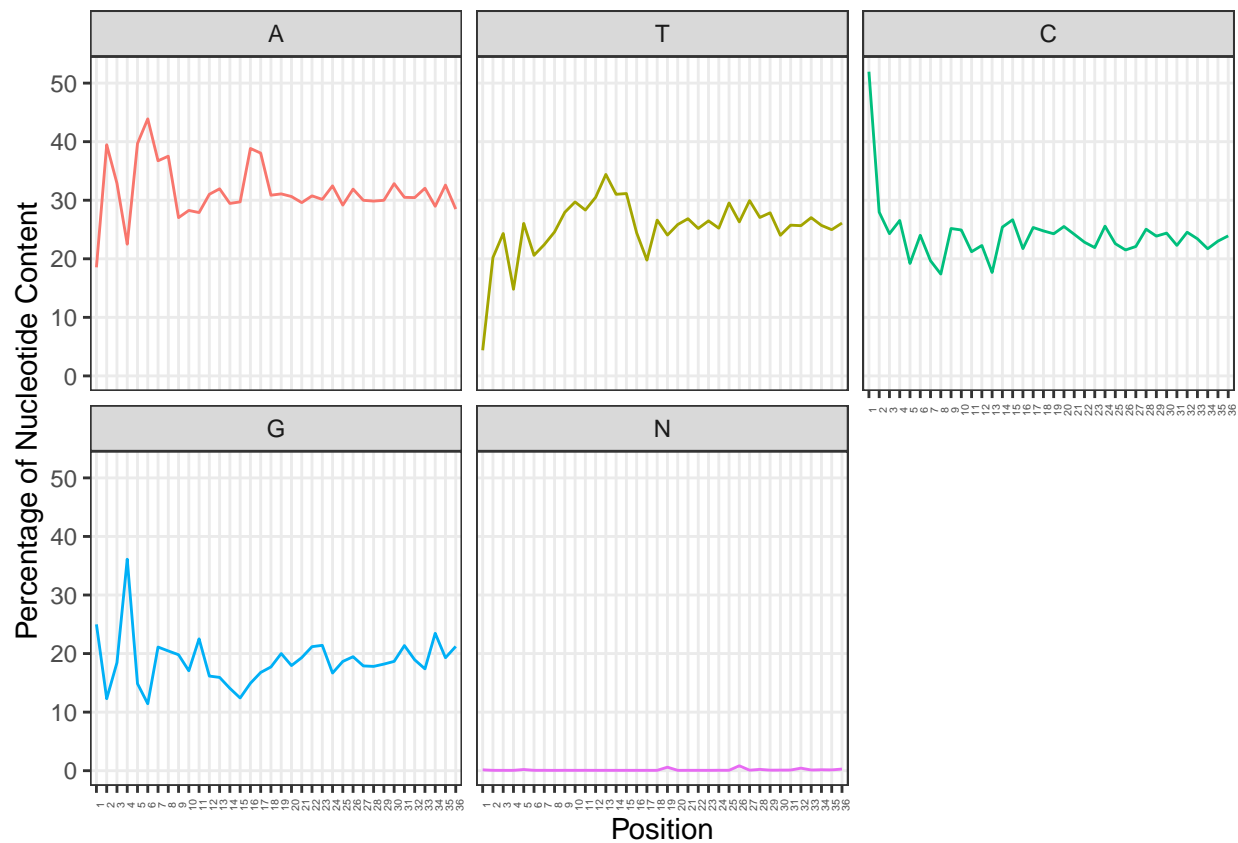
## Warning: package 'purrr' was built under R version 4.3.2

## Warning: package 'dplyr' was built under R version 4.3.2

## Warning: package 'stringr' was built under R version 4.3.2

## Warning: package 'lubridate' was built under R version 4.3.2

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    3.4.4      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```



## CODE

```
library(tidyverse)

### START QUESTION 1 CODE ###

## START QUESTION 1 PART B CODE ##

# Loading in summary statistic data
summary_stat_data <- read.delim("C:/Users/domin/Documents/Biostatistics Masters Program/Spring 2024/SMG

# First column name came in as X.column. Changing that
colnames(summary_stat_data)[1] <- "Column"

# Transforming data into long format
summary_stat_data_2 <- summary_stat_data %>%
  select(Column, A_Count, C_Count, G_Count, T_Count, N_Count) %>%
  pivot_longer(cols = ends_with("Count"),
               values_to = "Counts",
               names_to = "Nucleotide",
               names_pattern = "(.)_Count") %>%
  mutate(Column = factor(Column),
         Nucleotide = factor(Nucleotide, levels = c("A","T","C","G","N")))

# Making spaghetti plots
summary_plots <- summary_stat_data_2 %>%
  ggplot(aes(x = Column, y = (Counts/3614610)*100, group = Nucleotide,
            col = Nucleotide)) +
  geom_line() +
  facet_wrap(~ Nucleotide) +
  labs(x = "Position", y = "Percentage of Nucleotide Content") +
  theme_bw() +
  theme(legend.position = "none", panel.grid.minor = element_blank(),
        axis.text.x = element_text(angle = 90, size = 4))

summary_plots

## FINISH QUESTION 1 PART B CODE ##

### FINISH QUESTION 1 CODE ###
```