KNN Coding Practice with the tidyverse!

Question 1

We will continue to work with the `restaurants` data of Quiz 2!

Load the data with this code! Make sure to load the data with the following code:

```
library(tidyverse)
restaurants <-
    read_csv("https://www.dropbox.com/s/zzwv6hgiorckeuv/berkeley_restaurants.csv?dl=1")
```

part a

Add a column to `restaurants` called `photo_review_ratio` which divides the number of photos posted to Google by the number of reviews posted to Google.

Save the `photo_review_ratio` value for *`"Top Dog"`* and *`"Crepevine Restaurant"`* and into an object!

part b

Filter the `restaurants` data in two ways:

1. Include only `restaurants` in `"Southside"` or `"North Berkeley"`.
2. Remove our **testing** points, *`"Top Dog"`* and *`"Crepevine Restaurant"`*, from the datset.

We will consider this filtered dataset to be our *training set*.

part c

Add two new columns to the training set from **part b**: one which takes the absolute distance of `photo_review_ratio` from the value you recorded for *`"Top Dog"`*, and the other which does the same but for *`"Crepevine Restaurant"`*.

If you don't know how to find absolute value in R, look up a function that will do so!

part d

Using k-Nearest Neighbors with $k = 7$, predict the neighborhoods of *`"Top Dog"`* and *`"Crepevine Restaurant"`* (`"Southside"` or `"North Berkeley"`)!

Hints:

1. What `dplyr` commands would be useful to help you out here?
2. The `head()` function prints out the first few rows of a dataset; consult the help file for the `head()` function with `?` if you need to.

part e

Were your predictions correct?