STAT 243 PS3

Riv Jenkins

1 c)

In the "Best Practices for Scientific Computing" article, the authors claim that writing code in the highest level language possible is preferred even when the final product will need to be written in a lower-level language. They claim this is because higher level languages require fewer lines of code, and the program can later be rewritten in a lower-level language. I do not agree that this will be faster overall because although you may spend more time coding in a lower-level language if you start with that language than you would if you were simply transposing from a higher level language, I don't think the time saved by transposing would equal the amount of time necessary to write the program in the higher level language. There are other arguments in favor of coding in a high level first, and only switching to a low level if necessary, but I do not think time savings is one of them.

2 a)

```
fulltext = readLines("http://www.gutenberg.org/cache/epub/100/pg100.txt")
```

```
start_point = which(!is.na(str_extract(fulltext, "1603")))
end_point = which(!is.na(str_extract(fulltext, "We were dissever'd"))) + 2
all_plays = fulltext[start_point:end_point]
Play_ends = c(0, which(!is.na(str_extract(all_plays, "THE END"))))

Create_Play_Vec <- function(all_plays, Play_ends){
  char_vec = character(length(Play_ends) - 1)
  for(i in 1:(length(Play_ends)-1)){
    char_vec[i] = paste0(all_plays[Play_ends[i]:Play_ends[i+1]], collapse = ' ')
  }
  return(char_vec)
}

play_vec = Create_Play_Vec(all_plays, Play_ends)
substring(play_vec[1], 1, 55)
```

```
## [1] "1603  ALLS WELL THAT ENDS WELL  by William Shakespeare "
```

b)

```
length(which(!is.na(str_extract(all_plays, "THIS ELECTRONIC VERSION"))))
```

```
## [1] 215
```