Y, Z, Ryan Quigley
Section 2
Group X
MATH 267A

Problem Set 4

1. According to the help page for the function `scan()`, "a field is always delimited by an end-of-line marker unless it is quoted." Graphically, lines 1 and 2 end with the user-specified delimiter `,`, but R reads `,\n`. Thus it determines that there is an empty string between these two delimiters. Additionally, the order in which the empty strings appear in the resulting vector hints at the explanation above. The last wonder pet on line 1 is Tuck and the first wonder pet on line 2 is Ming-Ming; in the resulting vector, these strings are separated by an empty string. Similarly for Ollie and The Visitor.

2. When `{r} header = TRUE`, R expects the first line to contain the names of the variables that are associated with each column of data. The `read.table()` argument `check.names = TRUE` by default. When set to `TRUE`, this tells R to check the names of the variables in the data to ensure that they are syntactically valid variable names, i.e, name consists of letters, numbers and the dot or underline characters and starts with a letter or the dot not followed by a number [See `?read.table`]. If the names are not syntactically valid, the `make.names` function is called, which will prepend the character "X" if necessary [See `?make.names`]. This functionality can be overwritten by setting `check.names = FALSE`; however, this does not allow the user to select individual variables by name using the `$` operator.

```
caffeine.bad <- read.table("caffeine.txt", header = TRUE, check.names = FALSE)
head(caffeine.bad, n = 2)
```

```
##     0 100 200
## 1 242 248 246
## 2 245 246 248
```

```
caffeine.bad$`100`
```

```
## [1] 248 246
```

3.

c) Report the time that it takes to compute $\hat{\beta}$

d)

$$\begin{aligned}
\hat{\beta} &= (X^T X)^{-1} X^T y \\
(X^T X)\hat{\beta} &= (X^T X)(X^T X)^{-1} X^T y \\
(X^T X)\hat{\beta} &= I X^T y \\
(X^T X)\hat{\beta} &= X^T y
\end{aligned}$$