p8105_hw1_xy2404.Rmd

Annie Yu

9/20/2018

Problem 1

Create data frame with 1.numeric 2. logic 3. character and 4. factor vectors

```
Homework1_df = tibble(
  vec_numeric = runif(10, min =, max =5),
  vec_logic = (vec_numeric>2),
  vec_char = c("This", "is", "the", "first", "homework", "for", "the", "data", "Science", "class"),
  vec_factor = factor(c("T", "F", "T", "T", "T", "F", "T", "T", "T", "T"))
)
```

The mean of the variable containing absolute values is

```
mean(Homework1_df$vec_numeric)
```

```
## [1] 1.936645
```

```
mean(Homework1_df$vec_logic)
```

```
## [1] 0.3
```

```
mean(Homework1_df$vec_char)
```

```
## Warning in mean.default(Homework1_df$vec_char): argument is not numeric or
## logical: returning NA
```

```
## [1] NA
```

```
mean(Homework1_df$vec_factor)
```

```
## Warning in mean.default(Homework1_df$vec_factor): argument is not numeric
## or logical: returning NA
```

```
## [1] NA
```

It can be observed that the mean of numeric and logic variable can be derived, while that of character and factor variable failed. It is obvious that numeric vector have mean. The reason that logic variable have mean is that the value was stored as 0 (FALSE) and 1 (TRUE). However, as the character and factor variable are consist of strings, R can not convert them into string, hence NA was then returned.