Assignment 4

Maobin Guo

Part I

1. Create missging values

The following code is used to replace 30% age with missing value randomly.

```
df[head(sample(20), 0.3 * nrow(df)), ]$age <- NA
```

	number	diameter	age
1	1	12	125
2	9	11.7	
3	3	7.9	83
4	4	9	85
5	5	10.5	99
6	6	7.9	117
7	7	7.3	69
8	8	10.2	133
9	9	11.7	154
10	8	10.2	
11	16	8.5	
12	11	5.7	
13	13	10.3	114
14	14	12	147
15	15	9.2	122
16	16	8.5	106
17	17	7	82
18	18	10.7	88
19	1	12	
20	3	7.9	