Assignment

Question-1

```
1  Assignment:1
2  Your first assignment is to find a missing values in a ↩
       dataset, replace it with mean or median of that column↩
        or in case of non-numeric column, add "not-known" in ↩
       the missing cell.
3  Overall, your assignment should perform three steps:
4  a. Read the CSV dataset available for this assignment.
5  b. The dataset include "na", "NA" and "--" value, you need↩
        to clean that dataset, as mentioned above.
6  c. Store the new dataset into CSV format.
7
8  Your code should detect these values in any dataset.
9  Note: You would be interested to explore dplyr functions ↩
       such as mutate(...), transmute(...), replace(...).
10 Deadline: Oct 21, 2019.
```

Assignment

Question-2

```
1  The code to read the dataset is as follow:
2  survey <- read.table("http://www.andrew.cmu.edu/user/achoulde/94842/data/survey_data.csv", header=TRUE, sep=",")
3  Answer the following questions:
4  (a) How many survey respondents are from MISM or Other?
5  (b) What % of survey respondents are from PPM?
6  (c) Use $ notation to pull the OperatingSystem column from the survey data
7  (d) Do the same thing with [,] notation, referring to OperatingSystem by name
8  (e) Repeat part (d), this time referring to OperatingSystem by column number
9  Deadline: Oct 21, 2019.
```