

# Home Work Assignment 1

Elham

2023-02-02

Stat 123 Homework Assignment 1 Due Wednesday February 1st by 8:00pm

1. Twenty bottles of water were randomly selected from a large collection of bottles in a company's warehouse. These twenty bottles are referred to as the:

(C is Correct) [Points:0.5]

2. A journalist wanted to know how his constituents felt about a planned rezoning. He randomly selected 234 names from the city phone directory and conducted a phone survey.

(a) What is the population of interest? All people living in the region where the journalists want the opinion about the planned rezoning, something similar [Points:0.5]

(b) What is the sample of interest? The 234 names from the city phone directory [Points:0.5]

(c) What type of sampling was used? Simple Random Sample [Points:0.5]

3. You can see a sample of dataset below. Patient ID Cancer Type Number of Visits  
Status 123-87 Lung 3 Death 134-99 Brain 4 Survival 135-46 Breast 5 Death 124-76  
Kidney 3 Death 132-54 Lung 6 Survival

(a) What are the individuals in the data set? (Patient\_ID) [Points:0.5]

(b) What are the variables in the data set? (Cancer Type, Number of Visits, Status) [Points:1.0]

(c) Determine if each variable is categorical or numerical. (Number of Visits is Numerical/ Cancer Type and Status are Categorical Variables.) [Points:1.5]

4. A journalist is trying to determine the average age of the BC residents who have received fines for violating restrictions related to COVID-19. He solicits data from a local police department and records the following ages for the 6 tickets given out by this precinct.

13; 10; 25; 34; 15

(a) What is the parameter of interest? (The average age of the BC residents) [Points:1.0]

(b) Estimate the sample variance. Round your answer to 2 decimal places. [Points:1.0]

Note: Some of the students may considered population proportion for this question. They explained why the population proportion is unknown. Those students also

should get the mark because I asked them to tell me whether they can compute the population proportion for this question.

```
individuals<- c(13,10,25,34,15)
variance= round(var(individuals), digits=2)
variance

## [1] 98.3
```

(c) Estimate the sample mean. Round your answer to 2 decimal places. [Points:1.0]

```
individuals<- c(13,10,25,34,15)
variance= round(mean(individuals), digits=2)
variance

## [1] 19.4
```

5. Create the following data frame called Sample ID Name Age Vote 1 Juan 22 TRUE 2 Maria 15 FALSE 3 Mark 19 TRUE [Points:1.0]

```
Sample<- data.frame("ID"=c(1,2,3),
                    "Name"=c("Juan","Maria","Mark"),
                    "Age"=c(22,15,19),
                    "Vote"=c(TRUE, FALSE, TRUE))

Sample

##   ID  Name Age  Vote
## 1   1  Juan  22   TRUE
## 2   2 Maria  15  FALSE
## 3   3  Mark  19   TRUE
```

There is no problem if some of the students have ignored the ID column and have created the following dataset.

```
Sample<- data.frame("Name"=c("Juan","Maria","Mark"),
                    "Age"=c(22,15,19),
                    "Vote"=c(TRUE, FALSE, TRUE))

Sample

##   Name Age  Vote
## 1  Juan  22   TRUE
## 2 Maria  15  FALSE
## 3  Mark  19   TRUE
```

(a) Create a character vector called Age which contains the values from the second column of the Sample. [Points:1.0]

```
# First solution
Age<- c("22","15","19")
Age

## [1] "22" "15" "19"
```

```
# Second solution
Age<- c(22,15,19)
Age<- as.character(Age)
```

(c) Calculate the number of TRUE votes in third column. [Points:2.0]

```
VoteCounts<- length(Sample[Sample$Vote==TRUE])

# Students can also use Sample[,3], Sample[["Vote"]],...
VoteCounts

## [1] 2
```

(d) Calculate the mean of the ages in the second column. [Points:1.0]

```
MeanAge<- mean(Sample$Age)

# Students can also use Sample[["Age"]]or Sample[,2],...
MeanAge

## [1] 18.66667
```

4. Create a list called CourseName and mention the name of courses you have taken in this semester and then create the second list called CourseUnits with the unit of each course. [Points:1.0]

(a) Combine list CourseName and CourseUnits together and named it Courses. [Points:1.0]

(b) Change the values of CourseUnits to character. [Points:1.0]

```
CourseName<- list("MATH 204", "STAT 123", "STAT 261")

#Students can use different course names

CourseUnits<- list(1.5, 1.5, 1.5)

#Students can use different course units
CombineList<- append(CourseName, CourseUnits)

# Students also can use c(CourseName, CourseUnits)
CombineList

## [[1]]
## [1] "MATH 204"
##
## [[2]]
## [1] "STAT 123"
##
## [[3]]
## [1] "STAT 261"
##
## [[4]]
## [1] 1.5
##
```

```
## [[5]]
## [1] 1.5
##
## [[6]]
## [1] 1.5

CourseUnits<- as.character(CourseUnits)
CourseUnits

## [1] "1.5" "1.5" "1.5"
```