DATA 363: D1D2 Checkpoint

D1 – 1:

- Make sure that the paper clearly reports the limitations of the inference method used and all the possible sources of error. The paper should also report the validity of the data used and state any statistical and substantive assumptions made. It should also report the steps taken to preserve the integrity of the data.
- Make sure that the contributions and intellectual properties of others are acknowledged.

D1 – 2:

- We need to consider whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations before analyzing data from such a study. We also need to consider the treatment of all the research subjects when evaluating the sources of data.
- When contemplating whether to participate in an analysis of data from a particular source, we have to refuse to do so if participating could reasonably be interpreted as sanctioning a violation of the rights of the individuals who have provided their information.

D1 – 3:

- A person who is not a statistician has the ethical obligation to support sound statistical analysis and expose incompetent or corrupt statistical practices.
- They must also strive to protect the freedom and responsibilities of statistical practitioners who abide by the ethical guidelines presented

D1 – 4:

- Firstly, we need to recognize the difference between an honest mistake and misconduct
- We also need to know and understand the procedures related to the misconduct and follow the prescribed procedures for the same.
- Maintain confidentiality during the investigation but disclose results to honestly to appropriate parties once they are available.
- Finally, after the investigation, support the efforts of everyone involved in it and help them resume their careers in as normal a manner as possible.

D2Checkpoint

The probability of a woman being selected out of the given sample space is 1/3. Therefore, on average, there might be 8 women in a sample space of 24

```
population<-c(1:60)
(subjects1<-sample(population, 24))
```

```
##  [1] 24 39 54 18 37 31 20 19 32 41  8 48 38 10 55 52 15 51  6  9 57 40 59 13
```

```
(subjects2<-sample(population, 24))
```

```
##  [1]  6 26 58  7 57 28 38  9  2 47 59 41  5 11 52 45 55 25 27  1 31 36 37 53
```

```
(subjects3<-sample(population, 24))
```

```
##  [1] 56 12 22 35 30  6 38 21 14 11 15 49 43 44 47 55 59  9 52 48 60  2  5 25
```

```
(subjects4<-sample(population, 24))
```

```
##  [1] 15 13 55  3 46 56 30 47 38 32 58  2 54 36  6  1  5 51 45 23 27 40 42 16
```

```
(subjects5<-sample(population, 24))
```

```
##  [1] 28 32 57  6 26 39 37 59 36 44 53 24  4 35 46 13  5 22 23  8 47  3  2 10
```

```
(subjects6<-sample(population, 24))
```

```
##  [1] 12 52 57 31  9 53 18 15 33 42 26 32 11  1  5 10 39 37 25 40 56 28 27 13
```

```
(subjects7<-sample(population, 24))
```

```
##  [1] 30 26 47 57 41 60 54 35 31 29 58  8 15 40 34 16 24 39 38  9 17 43  7 42
```

```
(subjects8<-sample(population, 24))
```

```
##  [1] 25 19 20 15 28 33 29 54 39 26  1 55 22 59 10 21 17 53  4 46 57 37 42  8
```

```
(subjects9<-sample(population, 24))
```

```
##  [1] 33 37 45 15  2 10 41 43 23 59 47 25 14 18 32 46 36 13  5 12 17 54 60 58
```

```
(subjects10<-sample(population, 24))
```

```
##  [1]  8 35 42 36 17 38 55 29 15 49 26 48 32 25 18 11  4 58 51  1 12 21 34 60
```

```
women<-c(8, 7, 10, 10, 6, 8, 8, 5, 7, 8, 9)
mean(women)
```

```
## [1] 7.818182
```

```
sd(women)
```

```
## [1] 1.537412
```

This number lies quite close to the original estimate of approximately 8 women in our selected sample.

```
strat_men<-c(sample(population[1:20], 12))
strat_women<-c(sample(population[21:60], 12))
(strat = c(strat_women, strat_men))
```

```
##  [1] 40 22 43 52 25 49 21 60 33 35 53 23 10  5 14  4  1  8 11 12  2 16  9 17
```

Here we can see that out of our population size, we have managed to select an equal number of men and women despite the lower number of women compared to men. This shows that the study we are planning to use this data for is not biased with respect to gender.