

ST344u2140365Week1

2140365

2023-10-09

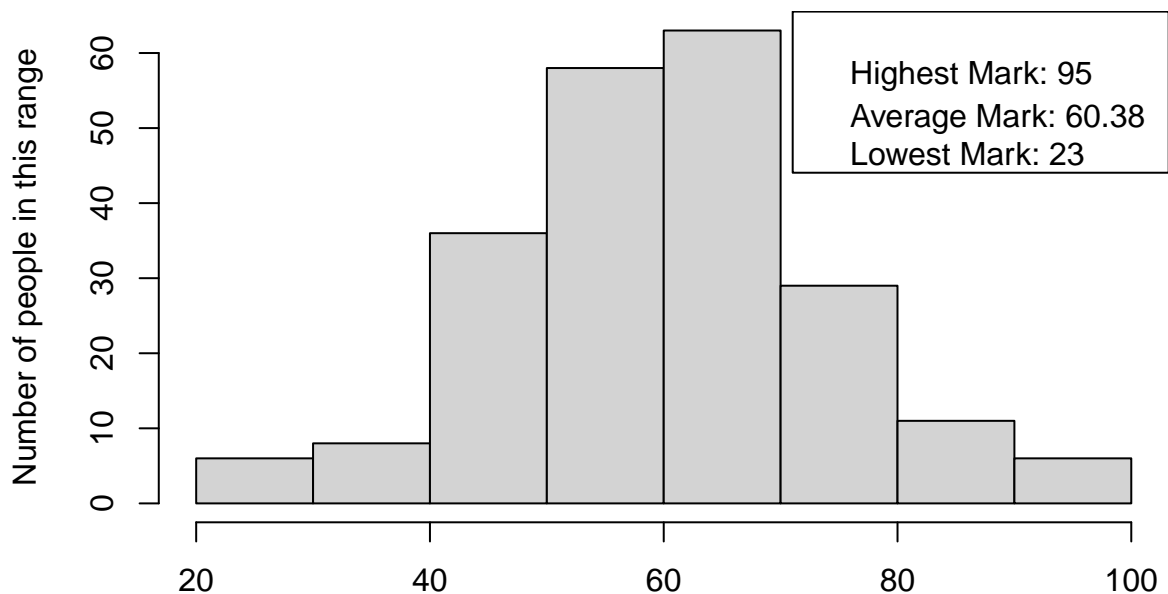
This is the first lab report of ST344 Professional Practice of Data Analysis. This report will contain 2 parts, First part is data analysis of four different human-made data; Second part is the analysis for Speed Data Study

PartI Initial Data Analysis

For PhysicsExam

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	23.00	51.00	61.00	60.38	69.00	95.00

Physics Exams Marks

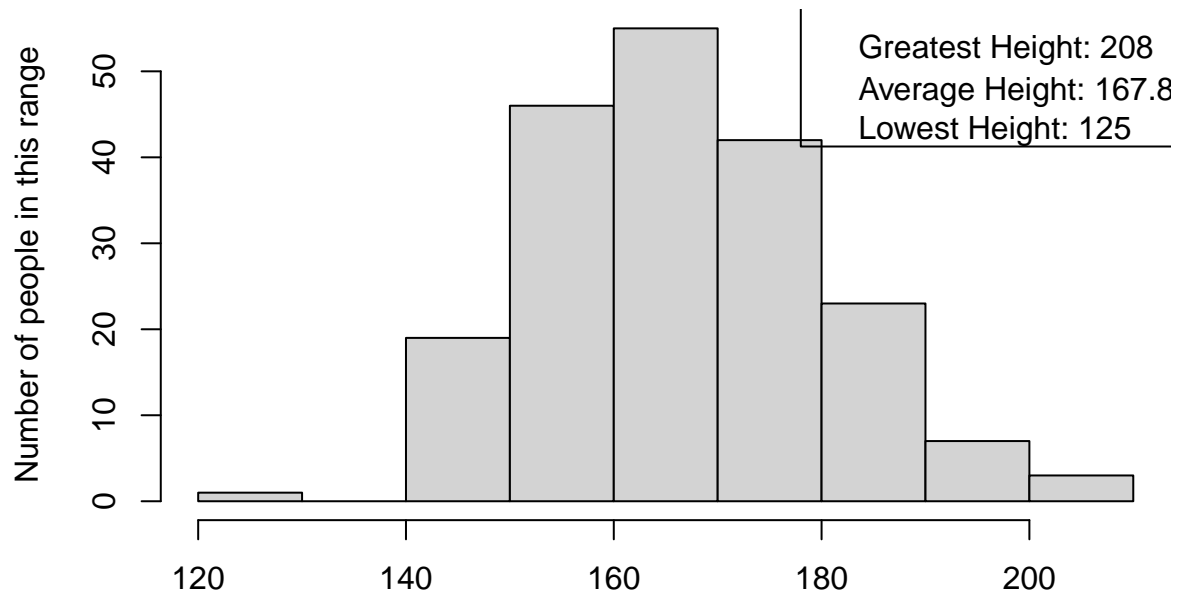


Many students scored between 60-70, potentially influencing the average of 60.38. Nevertheless, extreme high and low scores could impact the average. Eliminating these extreme values might enhance the accuracy of the average score.

For Height

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	125.0	157.0	168.0	167.9	178.0	208.0

Height of Cricket Players

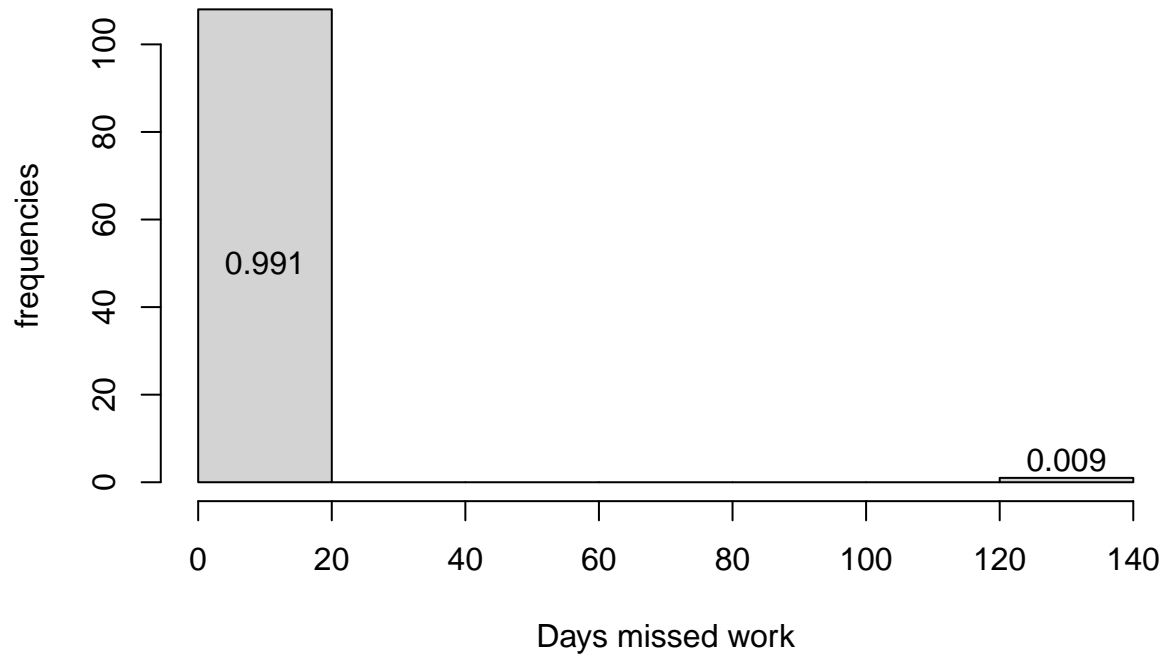


ers' heights range from 208cm (tallest) to 125cm (shortest), with most falling between 150cm to 170cm. Interestingly, height might not be a key factor in selecting cricket players.

For Nurse Days

```
## Days off with frequencies
## 0 1 2 3 4 5 19 129
## 67 27 7 1 1 4 1 1
```

Number of days work missed due to illness in 2022

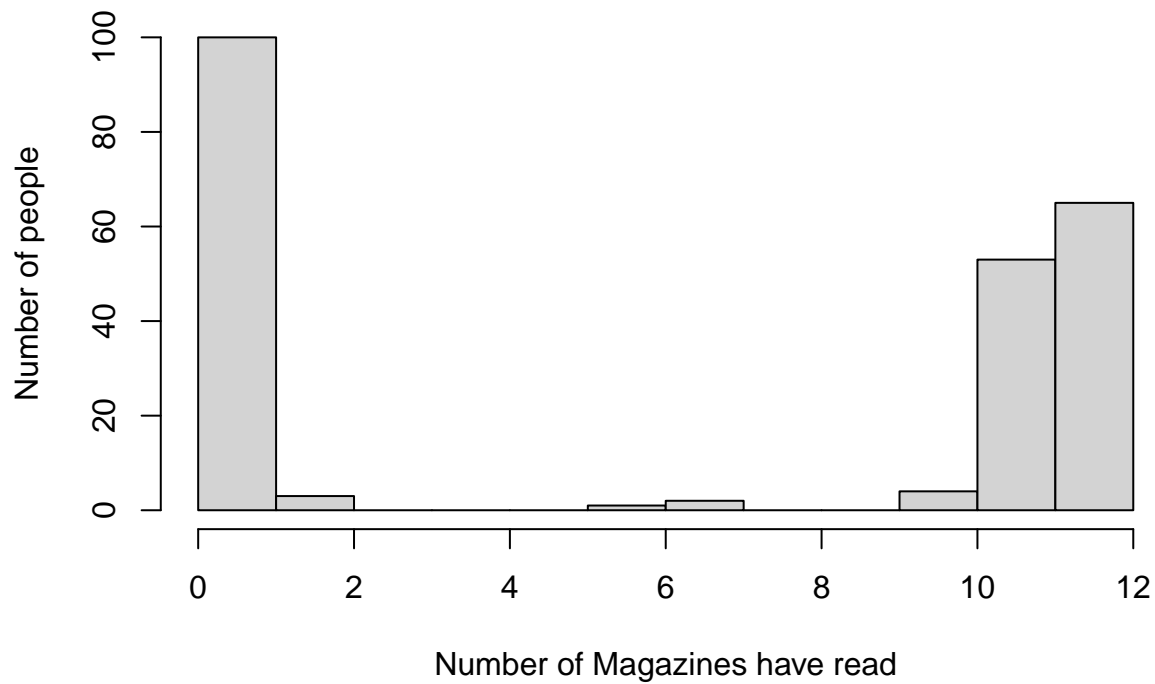


Almost all nurses missed work between 0-20 days, takes 99% of all. 67 people did not miss any workdays in 2022. Miss 120 workdays is quite unusual may be due to serious illness.

For Magazine

```
## Number of Magazines read and frequencies
## 0 1 2 6 7 10 11 12
## 43 57 3 1 2 4 53 65
```

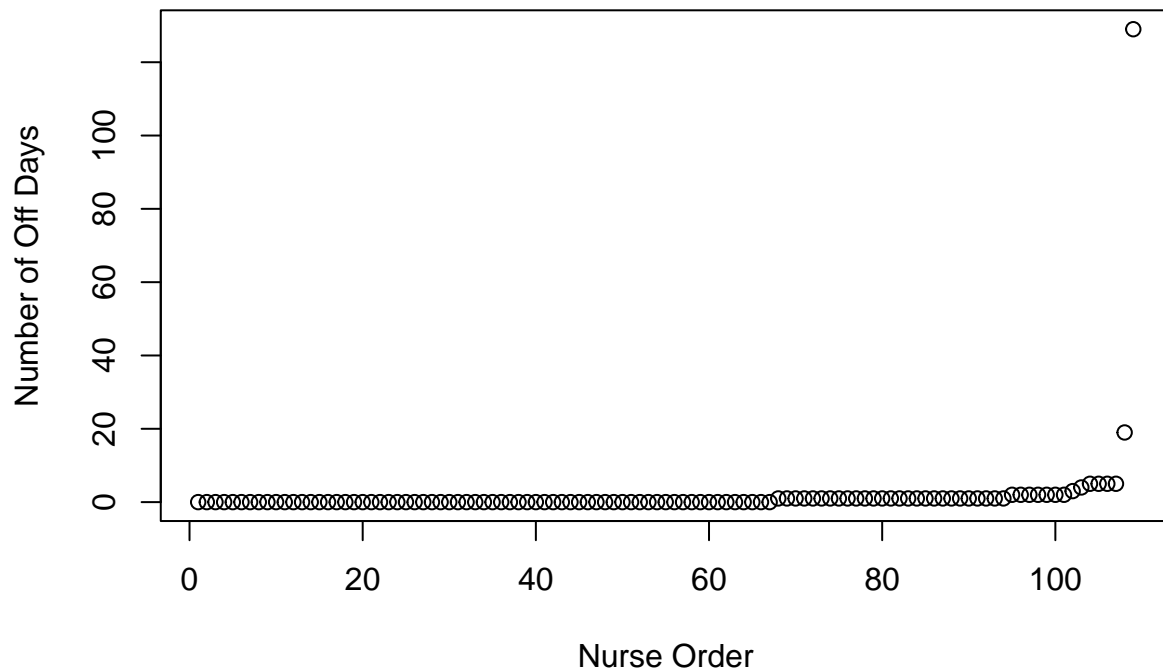
The number of issues of the magazine Football Monthly read in 202



There are many people who read 'Football Monthly' regularly and many who don't read it at all. It is rarely to see people read halfway and then stop.

Poor Data Nurse

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	0.000	0.000	1.982	1.000	129.000



It is

hard to define the meaning of Xlab, calculation of Mean and median makes no sense because of the extreme value.

PartII Speed Data

```
##  Gender, Average Age, NUmber of People, Population Percentage,
## 1   Male           26.6             277             0.5
## 2  Female           26.1             274             0.5
##  Most Popular Activity, Percentage of High Scores in Shared Interest,
## 1                sports             0.437
## 2                dining             0.549
##  Most common Activities Preferred by Highest Shared Interest Scorers
## 1                dining
## 2                dining
```

From the data, it's evident that male and female participants in Speed Dating share similar average ages, around 26 years old, with comparable numbers. However, males favour sports, while females prefer dining, influencing their shared interests. High-scoring males commonly enjoy dining, aligning with females' preferences. Top-scoring females still favour dining, potentially influenced by other factors. Moreover, the percentage of males scoring high in shared interests is lower, likely tied to their different activity preferences.
