

Homework 1 by Sana Khan

I produced the material below with assistance from our textbook

1. Define the following:

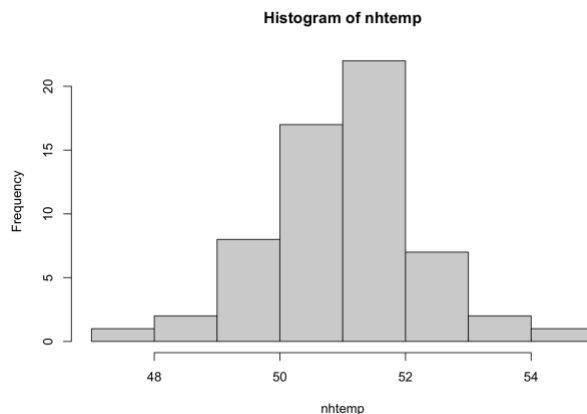
- Mean – the average for the values in your data set
- Median – the middle value for your data set
- Mode – the most common value occurring in your data set
- Variance – The average of the squared deviations from the mean. It's used to understand how disperse your data set is
- Standard deviation – The square root of the variance
- Histogram – A graph to show the number of observations plotted on an x and y axis
- Normal distribution – Follows a bell curve, which shows that the most common values are centered together
- Poisson distribution – plots how often observations happen between a specified time period

2. For this question I picked the Lake Huron dataset

```
>
> data()
> summary("LakeHuron")
      Length      Class      Mode 
      1 character character 
> mean(LakeHuron)
[1] 579.0041
> summary(LakeHuron)
      Min. 1st Qu.  Median     Mean 3rd Qu.     Max. 
576.0   578.1   579.1   579.0   579.9   581.9
```

-
- The mean is defined as the average of the dataset. In the Lake Huron data, the mean is 579, which means the average level of the lake was measured at 579 feet
- The Median is defined as the halfway value. In this data set the Median is 579.1, which means that the halfway value is 579.1

3. For this question I picked the nhtemp dataset.



- a. This follows a bell curve, which means it has a normal distribution. I believe this fits into a normal distribution because temperatures in this location would be similar year round.