Bhargavi Poyekar

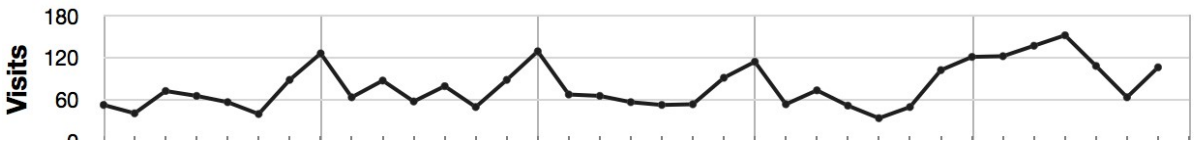
BE COMPS

2018130040

# Data Science, 2022
# Tut 8: Information Visualization

Question 1

The data shown here are the number of visits to a university website for a particular statistics course. There are 90 students in the class.




1. What are the names (type) of the 2 plots shown?

   1st = Time series plot

   2nd = Sparkline Plot- same data in compact form without axes

2. List any 2 interesting features in these data.
   a. No. of students in class are 90, but on few days, visits are greater than 90. So, some students might have visited the website more than once per day.
   b. Every 7th day, from 11jan to 25jan, there are sudden spikes, which suggests there might be some king of assignment deadline.
   c. In 1st week of February, a greater number of students have visited the website. I t might be because of upcoming exam.

Question 2

What are the names of the axes on a bar plot?

Usually, x-axis represents the different categories and y-axis represent the values.

Question 3

Which types of features can the human eye easily pick out of a time series plot?

The human eye easily pick out the trends and patterns, also the spikes and gaps.

Question 4

Why is the principle of minimizing “data ink” so important in an effective visualization? Give an scientific or engineering example of why this important.

The principle of minimizing “data ink” reduces the time or work to interpret that plot, by eliminating elements that are non-essential to the plot’s interpretation. It is important in situations that are time and safety critical, for example operation rooms in medical facilities, operator control rooms, etc.

Question 5

Describe what the main difference(s) between a bar chart and a histogram are.

| **Bar graph** | **Histogram** |
|---|---|
| Bar graph is the graphical representation of categorical data | Bar graph is the graphical representation of grouped data in continuous manner |
| There is equal space between each pair of consecutive bars. | There is no space between the consecutive bars |
| The height of the bars shows the frequency and the width gap is zero. | The frequency of the data is shown by the area of rectangular bars |