Research Experience for Teachers: Mitigating Natural Disasters

NDSU | NORTH DAKOTA STATE UNIVERSITY

Lesson #1 The Shape of Data

Subject Area	Mathematics
Grade Level	9-12
Prior knowledge	None
Time required	2 class periods

Summary

Courses in Introductory Statistics generally have two part: Descriptive statistics and Inferential Statistics. Descriptive statistics is concerned with displaying data and describing it in terms of *center* and *variability*. One of the first tasks after collecting raw data is to create a display. A histogram, dot plot or box plot gives the analyst a general sense of the data and its basic characteristics such as shape, center, and variability. In this lesson students will look at histograms of single variable data and learn to characterize the shape of data. They will then use technology to create data displays.

Education Standard

NCTM Principals and Standards

Select and use appropriate statistical methods to analyze data

- For univariate measurement data, be able to display the distribution, describe its shape, and select and calculate summary statistics.

I can

I can identify the shape of a probability distribution.

Introduction

Dr. Nic's Math and Stats video: *Random Variables and Probability Distributions*.
https://creativemaths.net/videos/video-distributions-1/

Notes and Examples

The shape of a distribution falls into two general categories: symmetric and skewed

Symmetric distributions can be further classified as

- Uniform
- Mound shaped
- U-shaped or bimodal

Skewed distributions are either left-tailed or right-tailed.

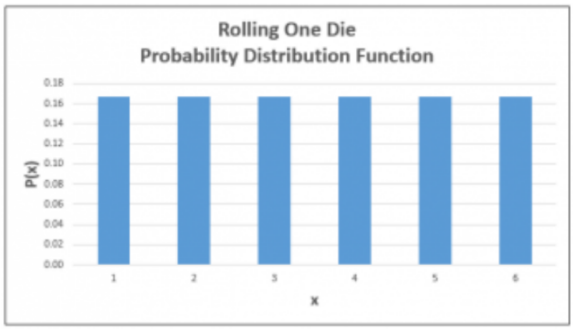
Symmetric Distributions

1. Uniform – All outcomes are equally likely.

This is an example of a probability distribution that results from rolling a single fair die.

x	1	2	3	4	5	6	sum
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{6}{6} = 1$

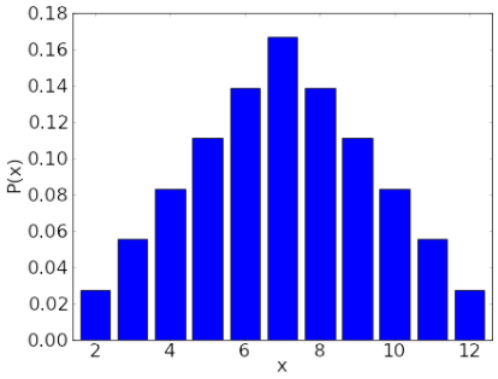
$$P(x) = \begin{cases} \frac{1}{6} & \text{if } x \in \{1,2,3,4,5,6\} \\ 0 & \text{otherwise.} \end{cases}$$



2. Mound shaped – outcomes near the middle are more likely.

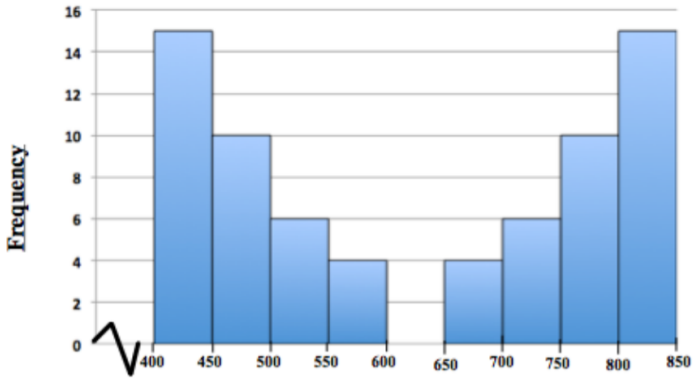
This is the sample space and a graph of the theoretical probability distribution from observing the sum of two dice

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12



3. U-shaped or Bimodal - The two peaks at either end make this distribution U-shaped. Because it has two peaks it can also be considered bimodal.

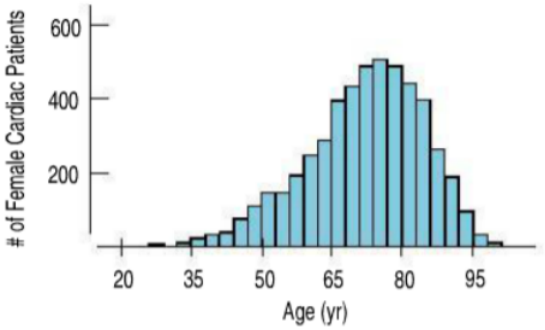
Score	Freq
400 - 449	15
450 - 499	10
500 - 549	6
550 - 599	4
600 – 649	0
650 – 699	4
700 - 749	6
750 - 799	10
800 - 849	15



Skewed Distribution

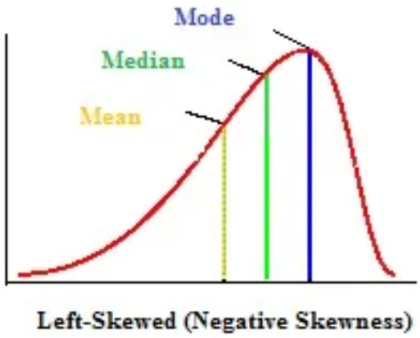
The thin end(s) of a distribution are call *tails*. If one tail stretches out further than the other tail, the distribution is considered ***skewed*** in the direction of the longer tail.

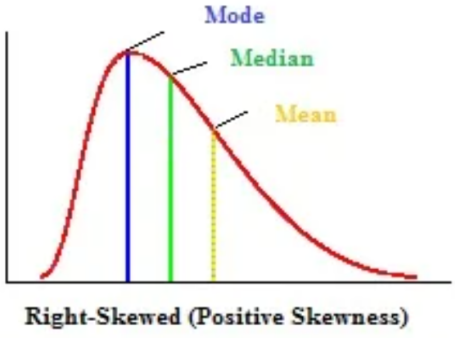
3. Skewed Left (negative skew or Left-tailed)



4. Skewed Right (positive skew or Right-tailed)

***To remember skewedness – look at your toes!

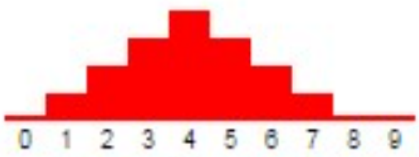


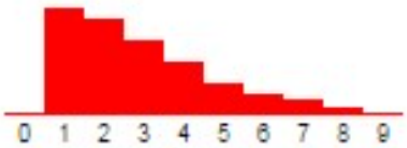


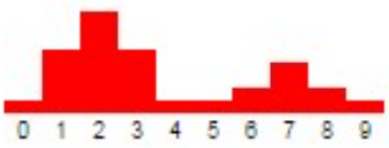
Formative Assessment

Describe the shape of the distribution using the suggestion in the word bank.

Curved	Symmetric	Slanty	Bi-modal
Skewed Left	Skewed Right	Uniform	Skewed Over
Mound Shaped	Crescent Shaped	Wiggly	Uni-modal

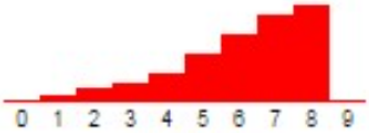




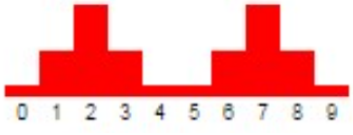


_____

_____



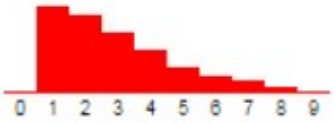
_____



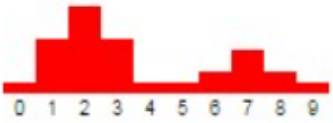
_____ _____ _____

Answers to formative assessment

Symmetric, unimodal, bell-shaped

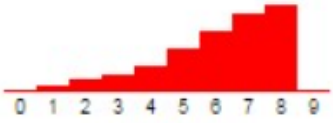


Skewed right



Non-symmetric, bimodal

Uniform



Skewed left

Symmetric, bimodal

Author:
Martha J. Nelson

Acknowledgements:
This curriculum was developed under the National Science Foundation RET grant #1953102. However, these contents do not necessarily represent the policies of the National Science Foundation, and you should not assume endorsement by the federal government.

Exercise #1 The Shape of Data

The Data:

The data in for this exercise was collected from a model of the intersection of Labeaux and County Road 18 in Albertville MN created using PTV VISSIM software. Videos of traffic simulation which generated the data:

0% AV saturation: https://www.youtube.com/watch?v=IkwWuVd-EAI
90% AV saturation: https://www.youtube.com/watch?v=S3Czr_taME0
3D view: https://www.youtube.com/watch?v=ylzQreGefqY

The Variables:

The software collected data on the following variables:

Average Vehicle Delay is the average number of seconds a vehicle is stopped at the intersection.

Average Queue Length is the average length in meters of the line of vehicles stopped.

Maximum Queue Length is the maximum length in meters of the line of vehicles stopped.

Number of Queue Stops is the number of stops made by all vehicles at the intersection.

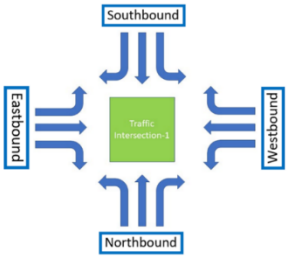
These variables are measured at each entrance to the intersection:

Westbound **(WB)**

Southbound **(SB)**

Eastbound **(EB)**

Northbound **(NB).**



Each simulation lasted 90 minutes (5400 seconds) with data collected every 15 minutes (900 seconds) and the first and last 15-minute intervals were discarded. Therefore, the data is collected for four time intervals for each run of the simulation:

900-1800 seconds, **1800-2700** seconds, **2700-3600** seconds, **3600-4500** seconds

Each variable is measured for seven autonomous vehicle saturation rates:

0% AV – all human driven cars

15% AV – 85% human driven cars

30% AV – 70% human driven cars

45% AV – 55% human driven cars

60% AV – 40% human driven cars

75% AV – 25% human driven cars

90% AV – 10% human driven cars

At each AV saturation rate, the simulation was run ten times. As a result, at each variable at each AV saturation rate there are 160 observations. One for each of the four directions at each of the four time intervals for each of the ten simulations.

At each saturation rate, the autonomous vehicles are tested using three different driving behaviors:

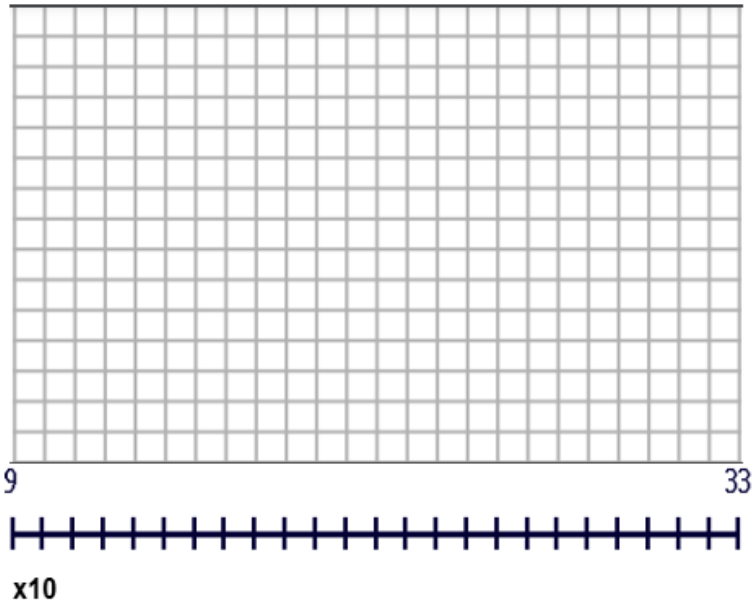
Cautious, **Normal**, **Aggressive**

	Queue stops								
	0% AV								
Run	TIMEINT	WB	QSTOPS	SB	QSTOPS	EB	QSTOPS	NB	QSTOPS
1	900-1800	1	123	2	312	3	113	4	274
1	1800-2700	1	104	2	299	3	100	4	279
1	2700-3600	1	127	2	234	3	124	4	297
1	3600-4500	1	157	2	247	3	142	4	228
2	900-1800	1	139	2	258	3	173	4	274
2	1800-2700	1	184	2	258	3	116	4	269
2	2700-3600	1	152	2	285	3	119	4	263
2	3600-4500	1	141	2	303	3	102	4	208
3	900-1800	1	117	2	271	3	134	4	252
3	1800-2700	1	116	2	291	3	117	4	235
3	2700-3600	1	107	2	249	3	134	4	280
3	3600-4500	1	126	2	233	3	110	4	207
4	900-1800	1	105	2	228	3	103	4	274
4	1800-2700	1	135	2	273	3	120	4	271
4	2700-3600	1	110	2	289	3	141	4	297
4	3600-4500	1	131	2	329	3	121	4	257
5	900-1800	1	176	2	267	3	122	4	301
5	1800-2700	1	138	2	248	3	131	4	262
5	2700-3600	1	167	2	280	3	128	4	278
5	3600-4500	1	141	2	312	3	139	4	289
6	900-1800	1	128	2	254	3	117	4	237
6	1800-2700	1	104	2	243	3	102	4	294
6	2700-3600	1	145	2	301	3	95	4	249
6	3600-4500	1	157	2	246	3	124	4	294
7	900-1800	1	107	2	311	3	139	4	280
7	1800-2700	1	148	2	262	3	135	4	320
7	2700-3600	1	113	2	289	3	102	4	253
7	3600-4500	1	132	2	275	3	108	4	267
8	900-1800	1	148	2	238	3	120	4	285
8	1800-2700	1	130	2	252	3	165	4	294
8	2700-3600	1	143	2	306	3	129	4	250
8	3600-4500	1	122	2	285	3	158	4	255
9	900-1800	1	141	2	240	3	150	4	246
9	1800-2700	1	133	2	264	3	144	4	275
9	2700-3600	1	140	2	269	3	145	4	328
9	3600-4500	1	143	2	233	3	120	4	256
10	900-1800	1	125	2	244	3	165	4	257
10	1800-2700	1	135	2	278	3	144	4	258
10	2700-3600	1	132	2	265	3	122	4	289

10	3600-4500	1	134	2	277	3	127	4	239

Using the number of queue stops data, make a histogram for each entrance to the intersection (WB, SB, EB, NB). For each display, describe the shape of the data.

Westbound:



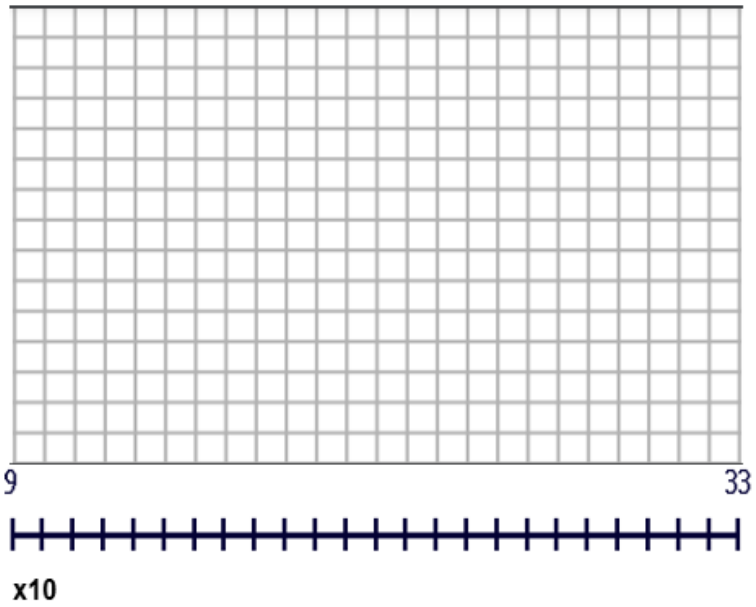
Shape:

_____

_____

_____

Southbound:



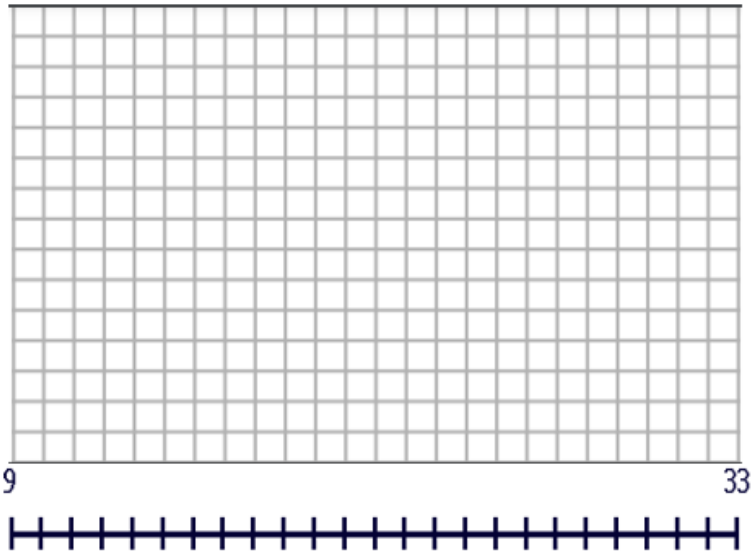
Shape:

_____

_____

_____

Eastbound:



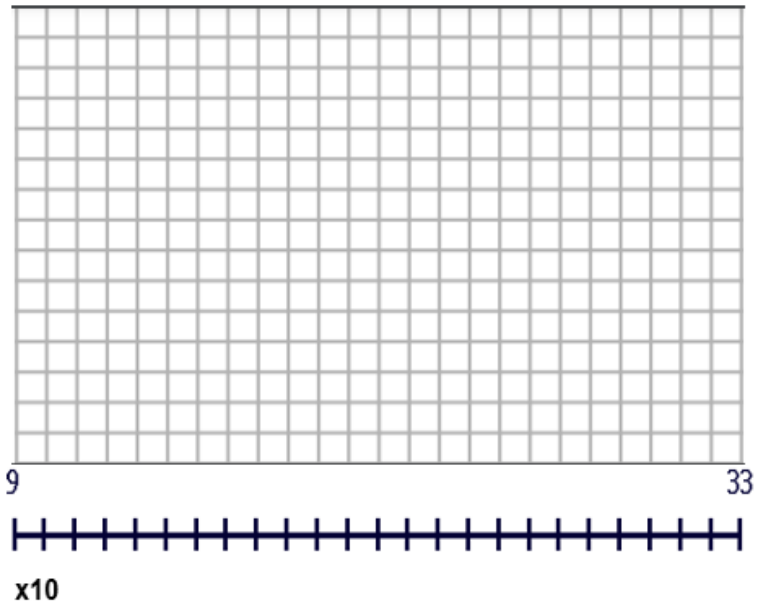
Shape:

_____

_____

_____

Northbound:



Shape:

_____

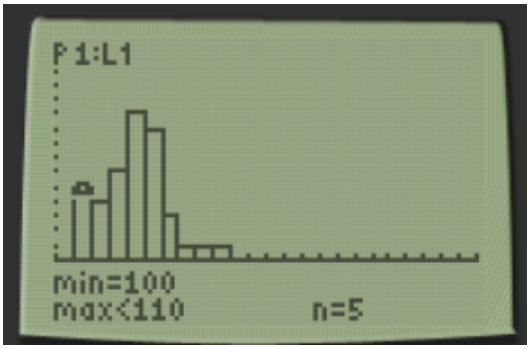
_____

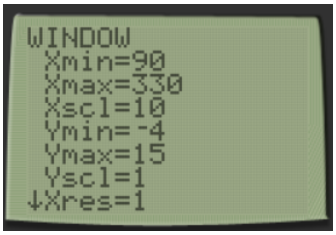
_____

Exercise#1 The Shape of Data Answers

Using the number of queue stops data, make a histogram for each entrance to the intersection (WB, SB, EB, NB). For each display, describe the shape of the data.

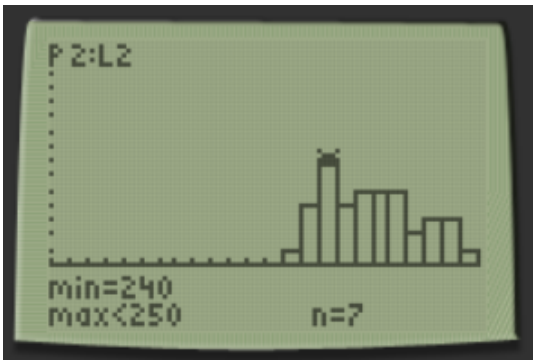
Westbound:





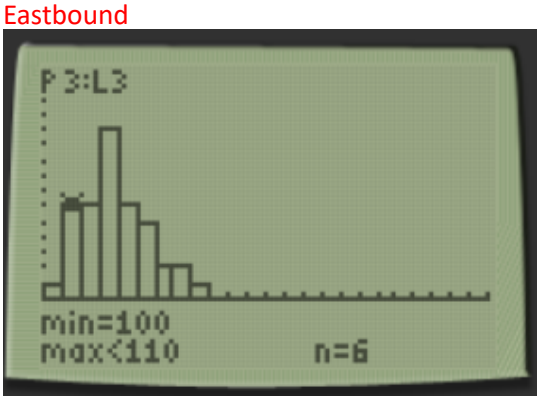
Mound shaped

Southbound:



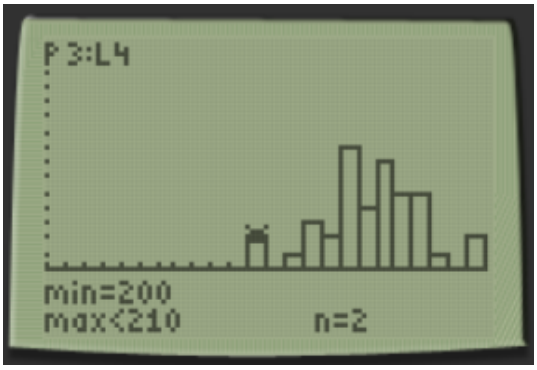
Mound shaped, approximately symmetric

Eastbound



Mound shaped, approximately symmetric

Northbound:



Mound shaped, approximately symmetric