

Statistics 177 Report

Data Visualization



Emily Mills

Fall 2019

Data Visualization

The following chart shows the raw data collected from 25 students at Middlesex Community College, Bedford Campus. Data was collected using convenience sampling. I asked 25 students who walked by in Henderson Hall on 9/18/2019 from 10:45am to about 11:30am. I was clear with subjects I was surveying that I was collecting data for a statistics course and asked if they were willing to answer a few questions. The data below represents self reported observations from 25 students regarding 4 different variables including favorite music styles, distance from school to home (in miles), commute time from school to home (minutes), and the number of credits taken this semester. Favorite music styles can be placed into the categorical variable category and produced qualitative data; Distance from school to home, commute time from school to home, and number of credits taken this semester can be placed into numerical variable category, and produced quantitative discrete data. Two students did not know the distance of their commutes and declined to answer.

Possible Bias:

1. I did not keep track of how many male vs. female students I surveyed. (Gender)
2. The time frame in which I surveyed students may have effected travel times to campus. (Students there for a 8am or 5pm class may encounter more traffic on their commute than a students there for a midday class.)
3. Personal bias - More likely to approach students who look friendly
4. Students seemed unsure of distance, a few declined to answer while many seemed to guess.

Analysis of Middlesex Community College Students-1

Students	Favorite Music Styles	School Distance From Home (Miles)	Commute Time to School from Home	Number of Credits Taken this Semester
1	Rap	7	18	9
2	Country	9	16	11
3	Rap	3	6	9
4	Rap	11	23	9
5	Rock	6	7	12
6	Pop	9	15	6
7	Country	8	18	3
8	Rap	Unknown	3	12
9	Pop	12	24	14
10	Rock	2	5	6
11	Country	18	42	14
12	EDM (electronic dance music)	11	25	9
13	Rock	10	20	6
14	Country	5	17	8
15	Rap	5	12	12
16	Country	2	5	6
17	Rock	5	8	6
18	Hip Hop	10	22	15
19	Rock	13	30	9
20	Country	Unknown	20	9
21	Rap	1	2	9
22	Pop	7	15	3
23	Rock	7	15	4
24	Hip Hop	10	21	3
25	Country	4	10	8

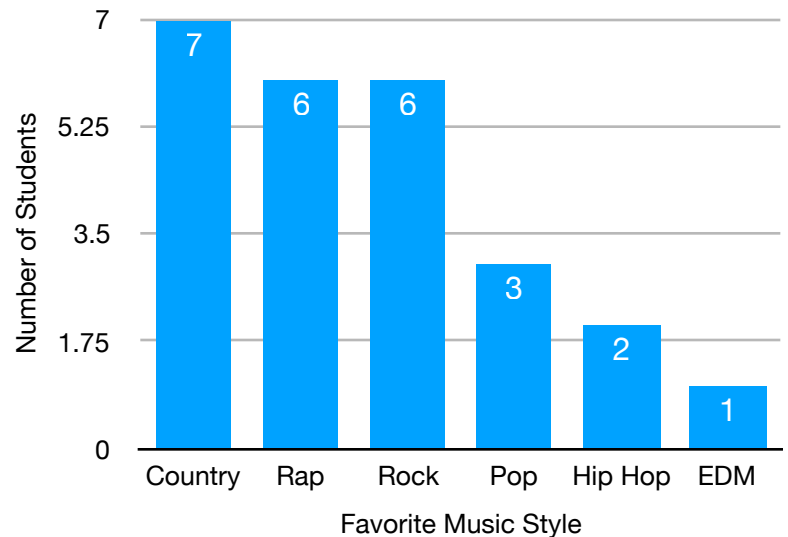
Favorite Music Styles Of Students

The raw data from the Favorite Music Styles of Students Spreadsheet was used to create the following bar graph. The graph simply displays the the qualitative data in a way that is easy for the reader to read and interpret. I chose a bar graph for those reasons as well as the fact that qualitative data is best presented in a bar graph or pie chart (according to lecture and textbook.). The Pie chart below also represents the same data in a different way.

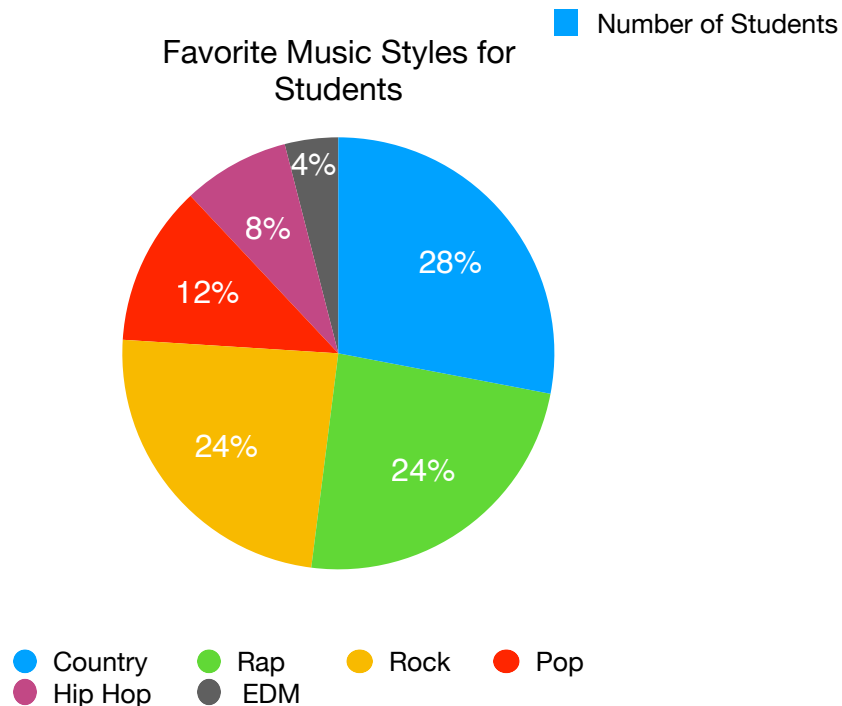
Favorite Music Styles of Students

Favorite Music Styles	Number of Students
Country	7
Rap	6
Rock	6
Pop	3
Hip Hop	2
EDM	1

Favorite Music Style of Students



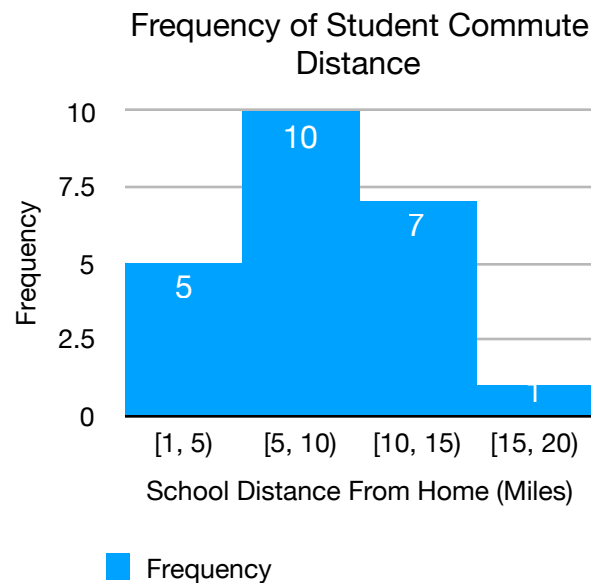
Favorite Music Styles for Students



School Distance From Home (miles)

I opted for a histogram to present the data for student commute distances to school. The histogram shows a bell shaped curve, but is missing two unknown values (two students did not know the distance.) However the histogram clearly presents the data to the reader, which will allow them to easily identify if there are any outliers, as well as see what the commute distance for most students in the sample is.

Students School	
School Distance From Home	Frequency
[1, 5)	5
[5, 10)	10
[10, 15)	7
[15, 20)	1
Unknown	2

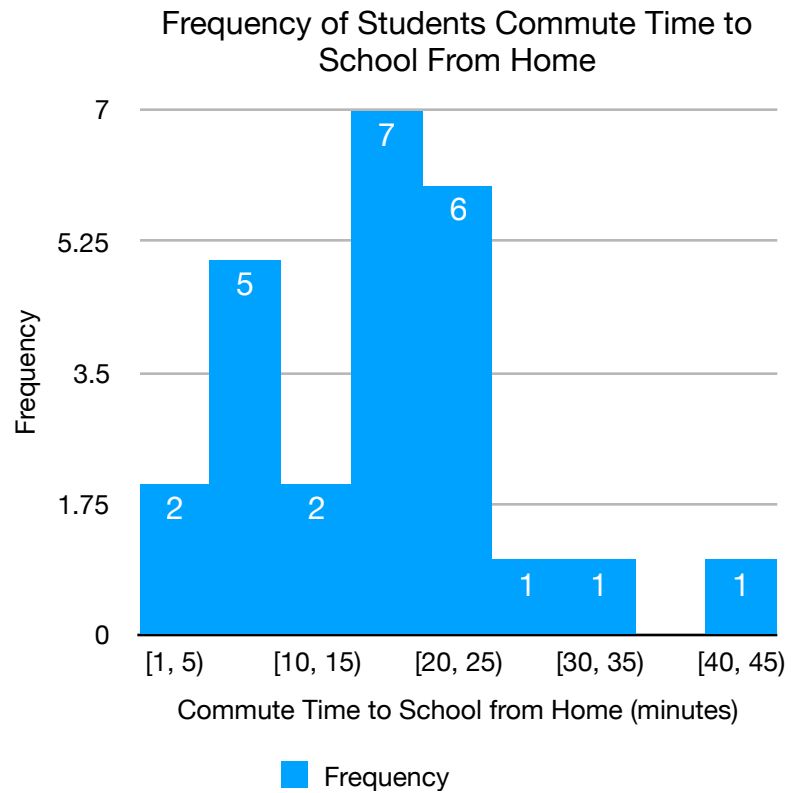


Students Commute Time from School to Home (Minutes)

I again opted for a histogram to present the data for student commute times to school. The histogram shows a slightly bell shaped curve. The histogram clearly presents the data to the reader, which will allow them to easily identify if there are any outliers, as well as see what the commute times for most students in the sample are.

Students Commute Time
From School to Home

Commute Time School to Home	Frequency
[1, 5)	2
[5, 10)	5
[10, 15)	2
[15, 20)	7
[20, 25)	6
[25, 30)	1
[30, 35)	1
[35, 40)	0
[40, 45)	1



Number of Credit Taken per Semester

I again opted for a histogram to present the data for the number of credits taken this semester. The histogram shows a bell shaped curve. The histogram clearly presents the data to the reader, which will allow them to easily make sense of the Data.

Frequency of Number of Credit Taken this Semester by Students

Number of Credits Taken This Semester	Frequency
[1,4)	3
[4, 8)	6
[8, 12)	10
[12, 16)	6

