AP Computer Science Principles
Ticket Out the Door
Set 1: Exploring One-Column Data

Name _____ Period _____

Skill 1.1 Exercise 1

For each data set below, identify the largest number and the smallest number.

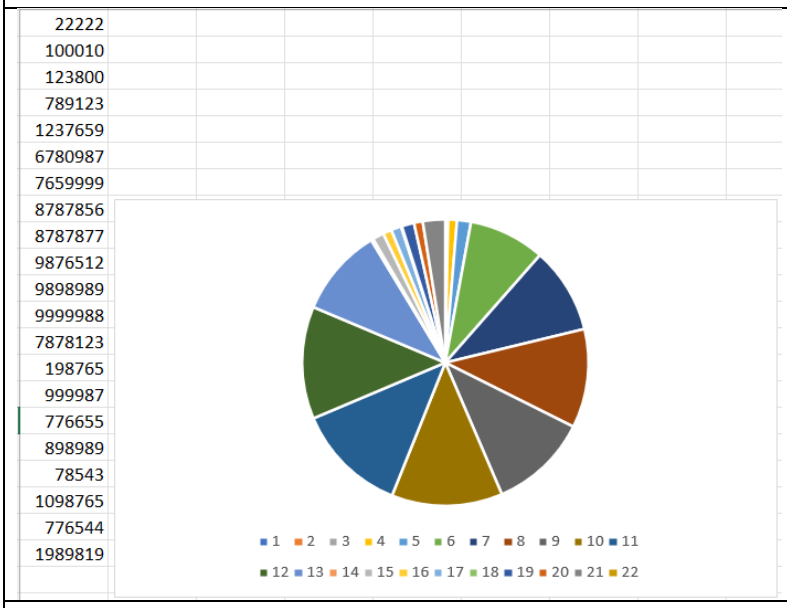
100010 636256 9898989 12121212 789123 123800 8907618

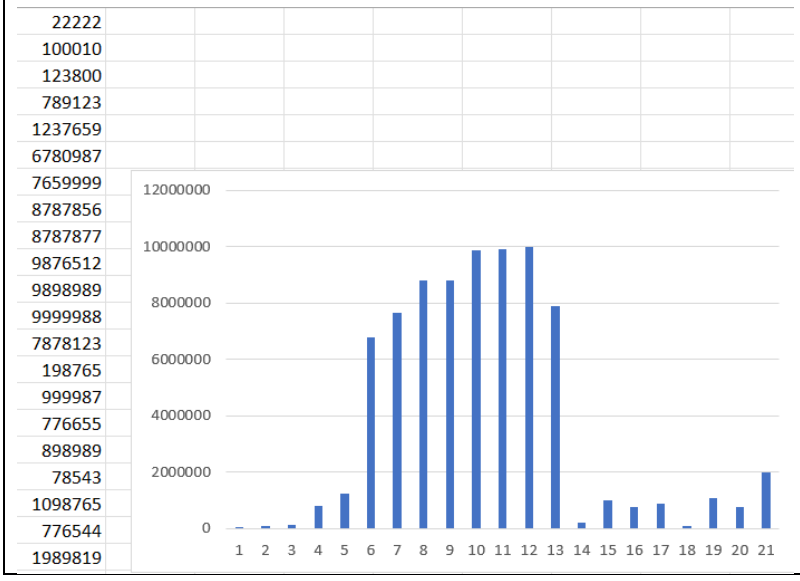
9875671 10101010 98111111 8787877 98989891

9876512 6780987 1237659 89898910 7424048 76654321

8787856 7659999 99887766 78956430 12914651

9999988 1628760 2222290 10926382 9187656 10987654





AP Computer Science Principles
Ticket Out the Door
Set 1: Exploring One-Column Data

Name _____ Period _____

Skill 1.1 Exercise 2
For which data set was it easiest to identify the largest and smallest number? Why?
Why do people make visualizations out of data?

Skill 1.2 Exercise 1
List the steps in the data analysis process.

Skill 1.3 Exercise 1

Below is a portion of a data set that contains information about different dog breeds. For each column, indicate whether it stores quantitative or qualitative data.

Breed Name	Breed Group	Bred For	Min Life	Max	Min Height	Max Height	Min Weight	Max Weight	Temperament
Affenpinscher	Toy	Small rodent hunting, lapdog	10	12	9	12	6	13	Stubborn, Curious, Playful, Adventurous, Active, Fun-loving
Afghan Hound	Hound	Coursing and hunting	10	13	25	27	50	60	Aloof, Clownish, Dignified, Independent, Happy
Airedale Terrier	Terrier	Badger, otter hunting	10	13	21	23	40	65	Outgoing, Friendly, Alert, Confident, Intelligent, Courageous

Column	**Data type (quantitative/qualitative)**
Breed Name	
Breed Group	
Bred For	
Min Life	
Max Life	

AP Computer Science Principles
Ticket Out the Door
Set 1: Exploring One-Column Data

Name _____ Period _____

Min Height	
Max Height	
Min Weight	
Max Weight	
Temperment	

Skill 1.4 Exercise 1

Open the dog data set, https://docs.google.com/spreadsheets/d/1dy2TrqRqXNcq-0k4ciLcATcPINv8u1eNPWhbJRuYYjU/edit?usp=sharing

For each column indicate whether the data would be best visualized using a bar chart, histogram chart, or neither.

Column	**Histogram, Bar, or Neither**
Name	
Breed Group	
Bred For	
Minimum Life Span	
Maximum Life Span	
Minimum Height	
Maximum Height	
Minimum Weight	
Maximum Weight	
Temperament	
Image	

For each column you indicated as best being visualized as a *histogram*, indicate a corresponding bucket size that would be suitable for visualizing the data.

Column	**Bucket Size**

AP Computer Science Principles
Ticket Out the Door
Set 1: Exploring One-Column Data

Name _____ Period _____

Skill 1.5 Exercise 1

The data below was collected from students using a survey. Indicate how you would *clean* the data by crossing out the incorrectly formatted data and replacing it with what was intended.

id	Age	Grade	FavoriteSubject	AverageHoursOfSleep	AverageHoursOfEntertainment
1	16	10	Math	7	2
2	15	10	Spanish	6	3
3	16	11	Spanish	7	two
4	17	10	CS	8	2
5	15	9	History	six	one
6	18	12	Computer Science	6	3
7	17	11	Bio	5	3
8	15	9	English	5	3
9	18	11	Music	8	0
10	15	11	Computer Science	8	3
11	15	11	Art	7	0
12	18	9	Computer Science	6	4
13	sixteen	nine	Art	6	1
14	15	9	Computer Science	6	0

You are a developer and have been asked to create a web-based form to collect the above data. What could you do to better ensure the user entered clean data?