STATS PROJECT: CAMILO AND JOE

METHODOLOGY

We employed a stratified sampling technique in which each stratum is a grade. There existed three pre-segmented groups (homerooms) within each stratum which we took to be representative of the stratum as a whole. We employed this technique so that we could make inferences about both our sub-populations as well as the generalized population. Our sample frame for each sub-population were proportional to the sub-population's (grade's) representation in the general population (High School Students). This was done in order to ensure that the ratio of the sample of each sub-population to the entire sample size was equal to the ratio of the sub-population to the general population. We used a random number generator to choose which homeroom to use to represent each sub-population. As the size of each of these homerooms did not always proportionally represent the sub-population, we chose the largest sub-population percent representation to be equal to the smallest randomly selected homeroom size and randomly omitted surveys taken from the other homerooms in order to maintain proportional representation. This was done to ensure that the sampling for each sub-population was limited to a single homeroom, eliminating the need to randomly select students from other homerooms to maintain proportional representation.

Population, Population Parameters, Sample Frame, Sampling Method

Our target population was Woods Charter School High School students.

As we only had one question that involved a quantitative response, we can only observe one population parameter that's not explicit in our raw data: average hours of sleep for the students of the high school. The average sleep the 12th graders received was 6 hours, 10th was 7 hours, and 9th was 8 hours. Thus, our predicted parameter for the entire high school population is 7 hours.

For our Sampling Frame we randomly selected one homerooms out of each grade. These were Mr. Durham's (9th Grade), Mr. Massengale's (10th Grade), Mr. Stewart's (11th Grade), and Mr. Temple's (12th Grade) homerooms.

We used a stratified sampling method in which each stratum was a grade with substrata being homerooms.

Sources of Bias

Potential sources of bias include:

- The population that was surveyed didn't take the survey very seriously, and it is likely that a good chunk of them circled answers at random.
- The assumption we made that homerooms are representative of their entire grade.
- Invalid answers from survey takers.
- Poor execution of methodology which led to the need to alter population size by excluding answers.
- Poor response rate in the 11th Grade homeroom which led to the 11th Grade being omitted completely.
- Errors due to rounding.
- Small sample populations meant that things like the 2% difference between Seniors and Sophomores were ignored because we were forced to only use 12 surveys from both grades (i.e. you can't exclude .34 of a person's data).
- The order in which question answers were listed especially for Question 1.
- We only counted responses that showed up more than one time for Question 2.
- Question 3 only asks about hours of sleep on one specific night, not over a long period of time.
- Some people, if not most, wouldn't take Question 5 seriously.

Raw Data

Question 1 - Choose your favorite genre of music

	Homeroom 2	Homeroom 5	Homeroom 8	Homeroom 10
Alternate	3	5	1	1
Classical		2		2
Country			1	
Electronic	1	1		
Hip Hop/Rap		2		3
Jazz				
Metal		1		
Pop	3	1		3
Rock	2	1	1	2
Soundtrack				
Other				
Invalid Answer	3	1	2	3

Question 2 - List your five favorite website

	Homeroom 2	Homeroom 5	Homeroom 8	Homeroom 10
gmail	11	10	4	9
facebook	11	12	5	12
youtube	11	10	2	9
google	8	5	4	4
tumblr		3		4
twitter	1	2		4
imdb	1			
game puma	1			
aq	1			
h&m			1	
modcloth			1	
anthropology			1	
al-jazeera			1	
npr			1	
wikipedia	5	4	1	1
national geographic			1	
google translate			1	
engrade	4	2	1	5
myxer				
drpepper				
pandora	1	3		3
netflix	1			
armorgames	1	1		
Invalid Answer		1	1	2
fmylife				1
thinkprogress				2
sixbillionsecrets				1

yahoo	1	2		
ffa				1
collegehumor	1			1
espn				2
intellicast		1		
4chan	1			1
420chan				1
steam powered		1		
newegg				
anandtech				1
urban		1		
banoiecollector				1
spanish dict.com		1		
easybib		1		
dictionary		1		1
howrse				1
commonapp				1
aol		1		1
cracked				
google-news		1		1
collegeboard		1		
kongregate		1		
mspaint		1		
oracle		1		
fanfiction		1		
att.net		1		
mugglenet.com		1		

Question 3 - How many hours of sleep did you get last night

Hours Of Sleep	Homeroom 2	Homeroom 5	Homeroom 8	Homeroom 10
1				1
2		1	1	
3				1
4		1		
5	1	1		2
6		2	2	3
7	4	1	2	5
8	4	5		1
9	2	2		1
10		1		
11	1			
WRONG ANSWER				

Question 4 - Which do you prefer?

Question 4	Homeroom 2	Homeroom 5	Homeroom 8	Homeroom 10
cake	10	8	3	12
pie	2	6	2	2

Invalid Answer				

Question 5 - Who do you prefer?

Question 5	Homeroom 2	Homeroom 5	Homeroom 8	Homeroom 10
michelle	11	4	3	2
palin	1	6	2	9
Invalid response		4		3

Analysis

For the purposes of analysis, we decided to omit Homeroom 8 (grade 11, Mr. Stewart's homeroom) because of the lack of received surveys (only 5). We accounted for this loss by entirely omitting the 11th grade from the population and thus as a stratum. We then recalculated each grades' representation in the population of 9th, 10th, and 12th graders and assigned weight according to this new representation (which hovered around 33% rather than 25%).

The graph of responses to Question 1 shows us that Country, Jazz, Soundtrack, and 'Other' are not popular at all among 9th, 10th, and 12th grade students at Woods. Metal and Electronic are also unpopular, while Alternative and Pop are extremely popular. Alternative and Electronic are

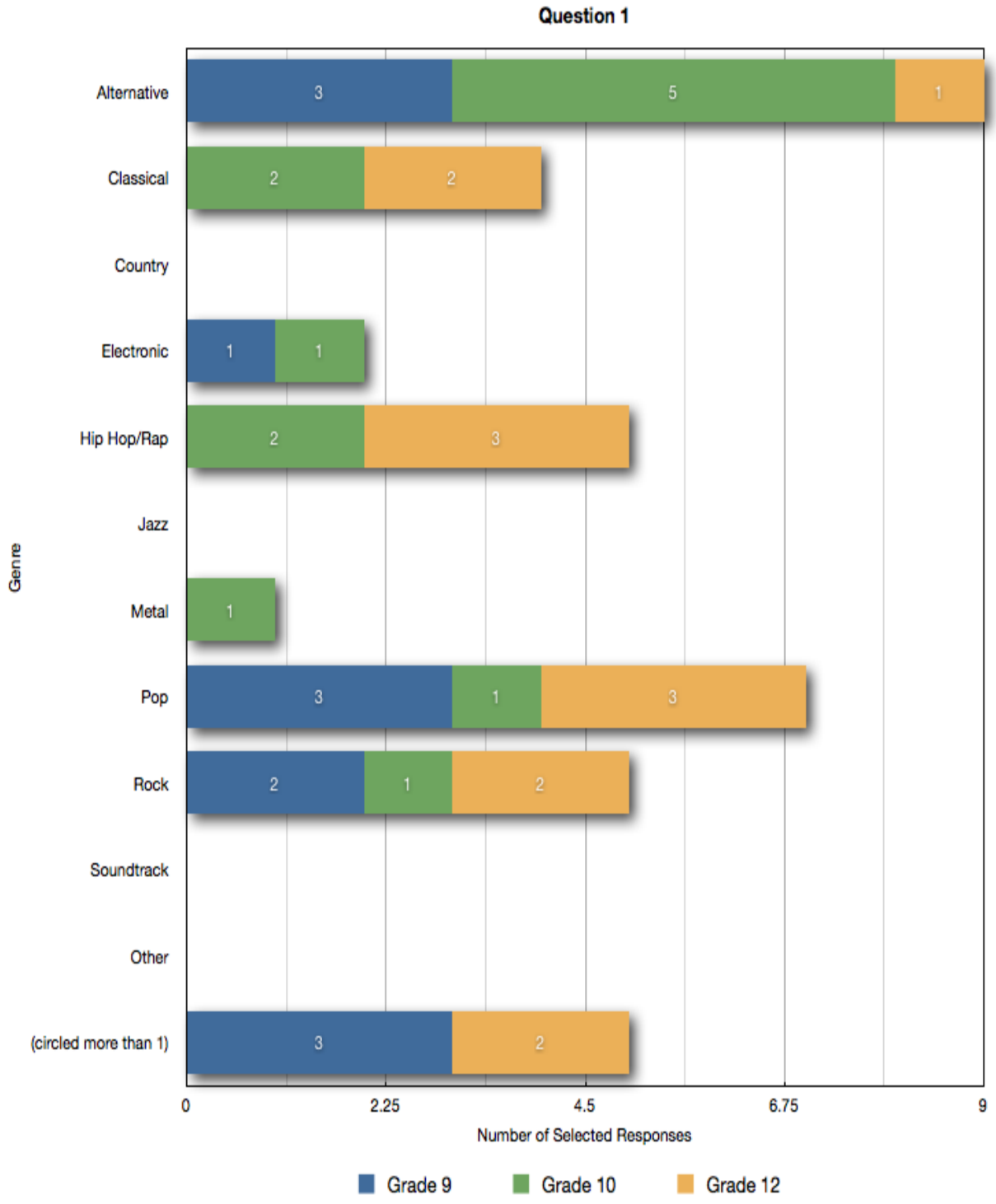
notably more popular among underclassmen these are the only genres where this is the case. Surprisingly, Classical, Hip Hop/Rap, and Rock are about equally popular. We can also see that a good amount of students are bad at following directions and chose more than one music type.

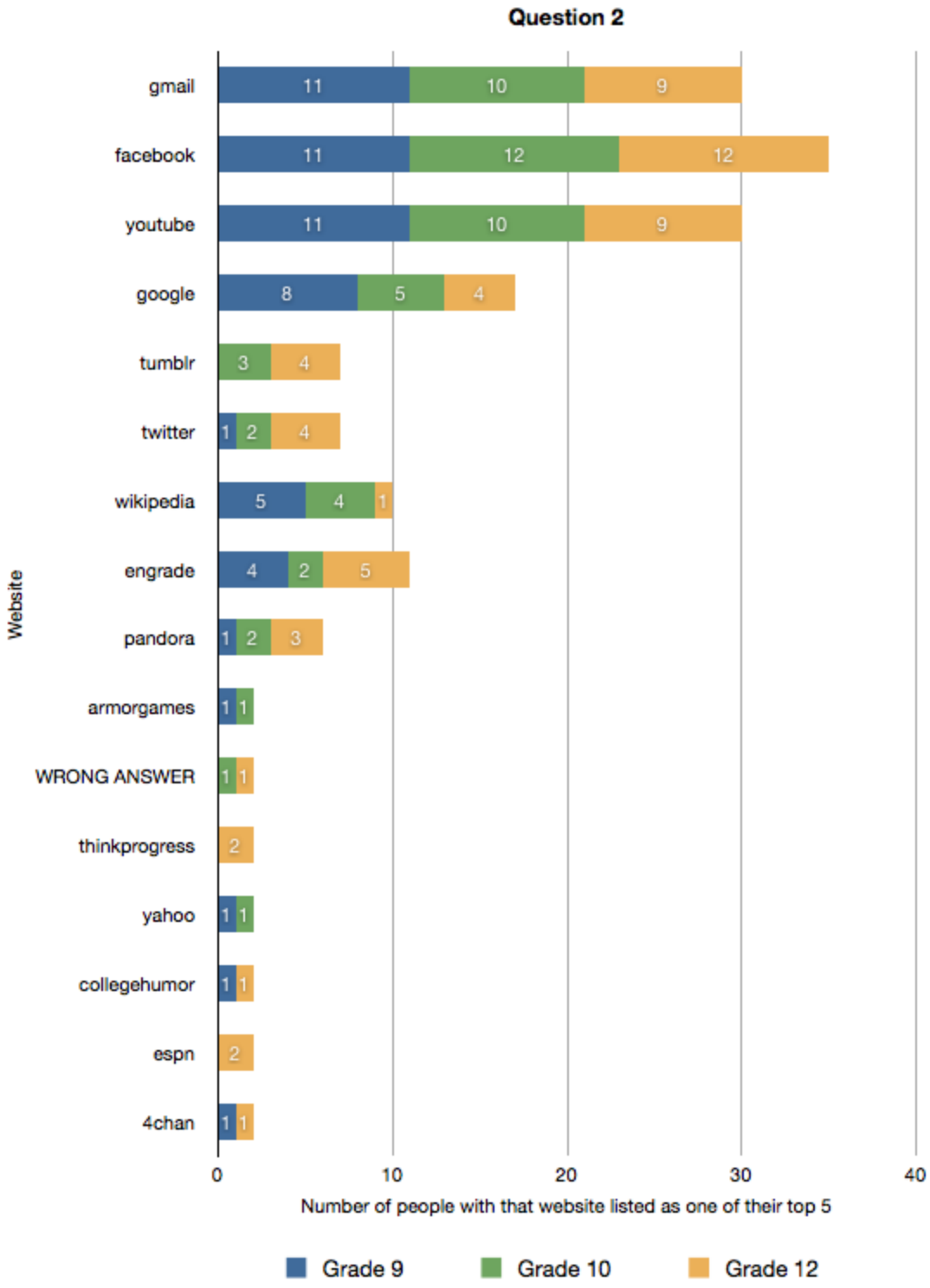
The graph of responses to Question 2 shows us that 9th, 10th, and 12th grade students at Woods are very active on Social and Entertainment sites on the internet with only four of seventeen responses not falling in either category (wikipedia, thinkprogress, engrade, and google) and 4 of the top 5 sites (gmail, facebook, youtube, and tumblr) falling in them.

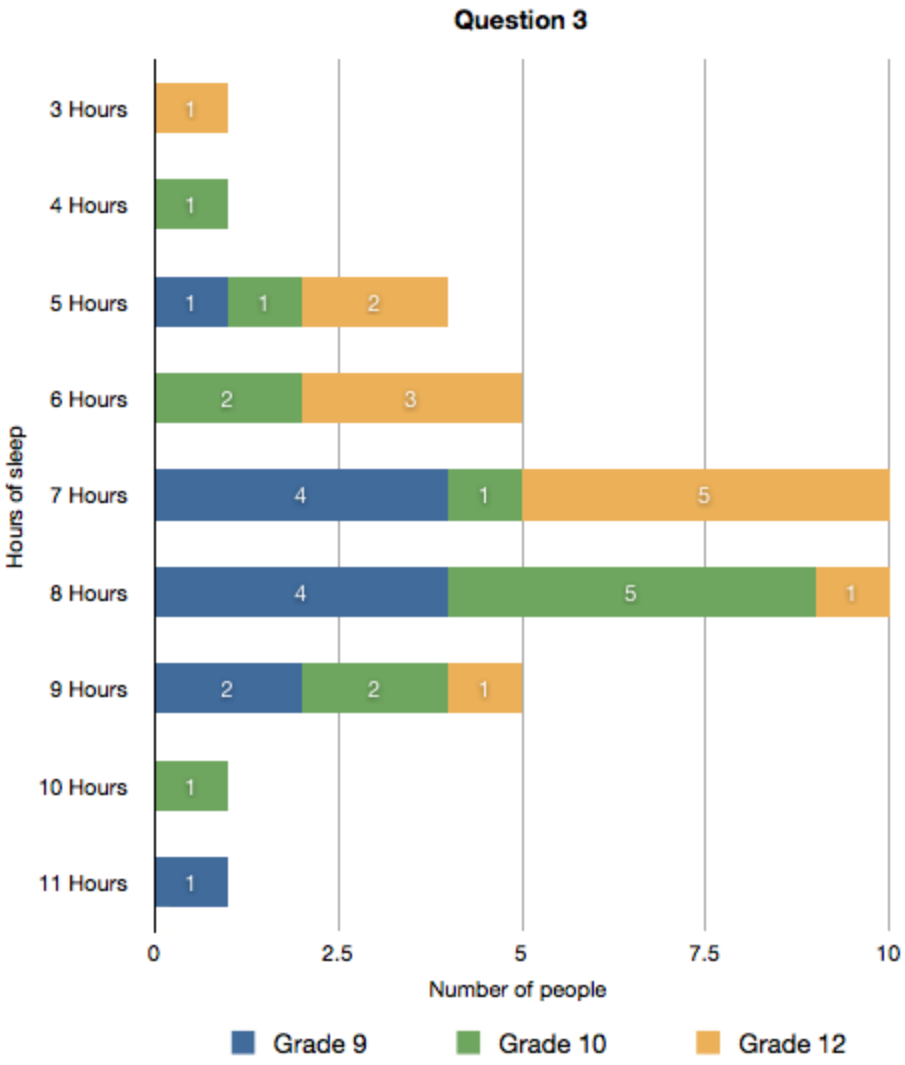
The graph of responses to Question 3 shows us that the vast majority of 9th, 10th, and 12th grade students at Woods got between 7-9 hours on October 27, 2011 (the day before the survey was done). A good chunk of them slept for 6 hours or less and a very small minority slept for 10 hours or more.

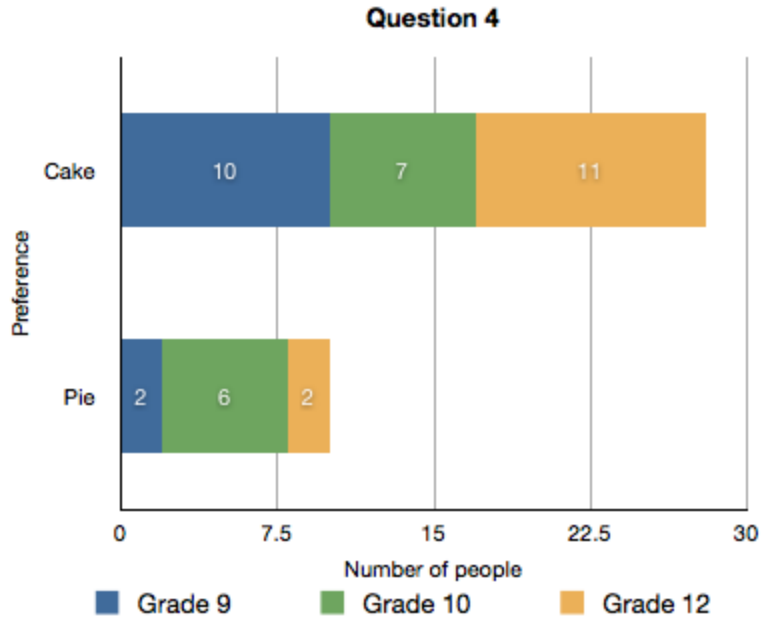
The graph of responses to Question 4 shows us that cakeis preferred over pie by 9th, 10th, and 12th grade students at Woods as a whole. However, in the 10th Grade cake is only slightly more popular than pie.

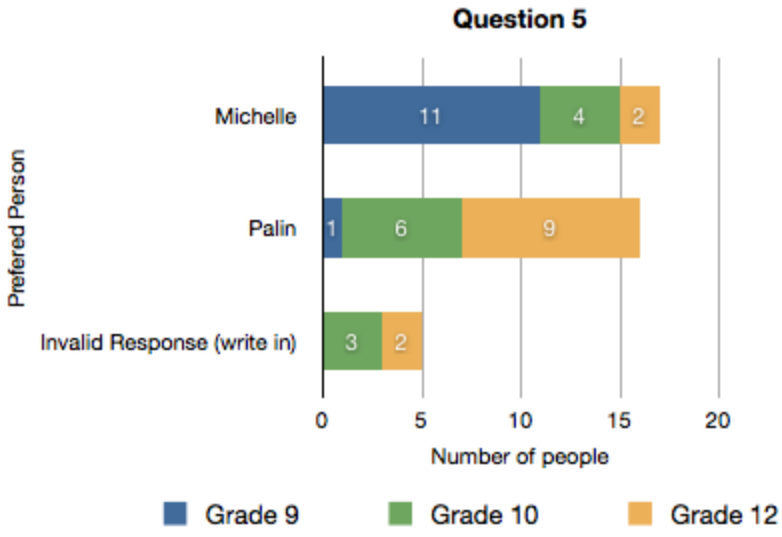
The graph of responses to Question 5 shows us that Michelle Bachman is largely preferred over Sarah Palin by 9th Grade students at Woods.12th Graders at Woods however, prefer Sarah Palin by as significant of a margin. 10th Graders at Woods are just about equally split. We can also see that about 23% of 10th Graders and about 15% of 12th Graders at Woods gave invalid responses while 0% of 9th Graders gave an invalid response.


Question 1
Alternative
Classical
Country
Electronic
Hip Hop/Rap
Jazz
Metal
Pop
Rock
Soundtrack
Other
(circled more than 1)
Genre
3
5
1
2
2
1
1
2
3
1
3
2
1
2
3
2
0
2.25
4.5
6.75
9
Number of Selected Responses
Grade 9
Grade 10
Grade 12


Question 2
gmail
facebook
youtube
google
tumblr
twitter
wikipedia
engrade
pandora
armorgames
WRONG ANSWER
thinkprogress
yahoo
collegehumor
espn
4chan
11
10
9
11
12
12
11
10
9
8
5
4
3
4
1
2
4
5
4
1
4
2
5
1
2
3
1
1
1
1
2
1
1
1
1
2
1
1
Website
0
10
20
30
40
Number of people with that website listed as one of their top 5
Grade 9
Grade 10
Grade 12


Question 3
3 Hours
4 Hours
5 Hours
6 Hours
7 Hours
8 Hours
9 Hours
10 Hours
11 Hours
Hours of sleep
1
1
1
1
2
2
3
4
1
5
4
5
1
2
2
1
1
1
0
2.5
5
7.5
10
Number of people
Grade 9
Grade 10
Grade 12


Question 4
Cake
Pie
Preference
10
7
11
2
6
2
0
7.5
15
22.5
30
Number of people
Grade 9
Grade 10
Grade 12


Question 5
Michelle
Palin
Invalid Response (write in)
Prefered Person
11
4
2
1
6
9
3
2
0
5
10
15
20
Number of people
Grade 9
Grade 10
Grade 12

CHECKLIST:

Population Parameters: A numerically valued attribute of a model for a population. We rarely expect to know the true value of a population parameter, but we do hope to estimate it from sampled. For example, the mean income of all employed people in the country is a population parameter.

SURVEY QUESTIONS:

1. What genre of music do you listen to the most?
    a. Alternative
    b. Classical
    c. Country
    d. Electronic
    e. Hip Hop/Rap
    f. Jazz
    g. Metal
    h. Pop
    i. Rock
    j. Soundtrack
    k. Other
2. Top 5 websites you visit?
3. How many hours of sleep did you get last night?
4. Cake or pie?
5. Michelle Bachman or Sarah Palin?

Perfect Population

This would've been our population had we executed our methodology perfectly.

Grade 09: 46 people	25.1366	TARGET: 25%
Grade 10: 47 people	25.6831	TARGET: 26%
Grade 11: 41 people	22.4044	TARGET: 22%
Grade 12: 49 people	26.7759	TARGET: 27%
Total: 183 people		

Homerooms chosen:

Homeroom 2 (Mr. Durham)	16 9th Graders (exclude 3)
Homeroom 5: (Mr. Massengale)	15 10th Graders (exclude 1)
Homeroom 8: (Mr. Stewart)	14 11th Graders (exclude 3)
Homeroon 10: (Mr. Temple)	18 12th Graders (exclude 4)
	63 Total, 52 after excluding

Total in each homeroom.....after exclusions

18 seniors ...14
14 11th graders ...11
15 10th graders ...14
16 9th graders. ...13

Total sample size = 63....52.