**CIS 362 Project 1**

**Shower Survey - Phase 2**

**March 10, 2020**

**Team 9 -** Jehu Barbee, Eric Botelho, Ashley Famlaro, Grady Landers, Michael Medeiros

**Introduction**

The survey conducted found data regarding the shower habits of UMass Dartmouth students. Attributes acquired from the focus group include gender, age, school year, residence status, average showers taken per week, and average shower length, as well as asking whether the student uses the gym or plays sports, showers outside of their residence, if they ever shower for longer than average and if they think they can reduce their shower time by one minute. Below is a table outlining the statistical findings of the collected data.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Attribute** | **Mean** | **Median** | **Std Dev** | **Variance** | **Range** | **Skew** | **Quartiles** | **Outliers** |
| Gender | 0.516 | 1 | 0.5 | 0.25 | 1 | N/A | N/A | N/A |
| Age | 21.112 | 20 | 5.822 | 33.896 | 51 | 6.14 | Q1: 19  Q2: 20  Q3: 21 | 69 |
| School Year | 2.227 | 2 | 0.936 | 0.876 | 4 | 0.071 | Q1: 1  Q2: 2  Q3: 3 | N/A |
| Resident | 0.51 | 1 | 0.5 | 0.25 | 1 | -.041 | Q1: 0  Q2: 1  Q3: 1 | N/A |
| Gym/Sports | 0.459 | 0 | 0.498 | 0.248 | 1 | 0.164 | Q1: 0  Q2: 0  Q3: 1 | N/A |
| Shower other places | 0.082 | 0 | 0.274 | 0.075 | 1 | 3.056 | Q1: 0  Q2: 0  Q3: 0 | 1, 1, 1, 1, 1, 1, 1, 1 |
| Avg. Showers per week | 6.612 | 7 | 2.735 | 7.482 | 16 | 1.034 | Q1: 5  Q2: 7  Q3: 7 | 15, 17 |
| Avg. Shower time | 16.031 | 15 | 9.238 | 85.336 | 58 | 1.739 | Q1: 10  Q2: 15  Q3: 20 | 60, 45 |
| Occasionally longer | 0.847 | 1 | 0.36 | 0.13 | 1 | -1.927 | Q1: 1  Q2: 1  Q3: 1 | N/A |
| Can reduce | 0.827 | 1 | 0.379 | 0.143 | 1 | -1.725 | Q1: 1  Q2: 1  Q3: 1 | N/A |

**Gender**

The values for gender were coded in such a way that 0 represents female, 1 represents male, and 4 people chose not to answer which is why skewness and quartiles could not be calculated. The mean value is 0.516, which in this case means 51.6% of the focus group identify as male. The median value is 1, which indicates that the value in the center of the data is 1, supporting that 51.6% of the sample space is male. The standard deviation and variance are 0.5 and 0.25 respectively, which are low to be expected because the only values are 0 and 1, which explains the range of 1.

**Age**

The mean value for age is 21.112 years old which indicates that the majority of the focus group were around 21 years of age. The median is 20 years old, expectedly close to the mean. The standard deviation and variance are 5.822 and 33.896 respectively, relatively low showing that in general the ages were not too far off from each other. The skewness is 6.14, somewhat high but can be explained by the outlier of 69, the max for this data that results in a range of 51. The first, second and third quartiles are 19, 20, and 21, meaning that the first 25% of students are 19 or younger, 50% are 20 or older, and the top 25% are older than 21, supporting the mean age of 21.

**School Year**

School year was coded from 1 to 5, 1 being freshman and 5 being any graduate student over 4 years into school. The mean value is 2.224 and median is 2, indicating that most of our focus group are sophomores. The standard deviation and variance are 0.936 and 0.868 respectively, reasonably low considering the data was coded to have a range of 4, which also explains the low skewness of 0.071. The first, second and third quartiles are 1, 20, and 3, meaning that the first 25% of students are freshman, 50% are sophomores, and the top 25% are juniors or higher. Expectedly there are no outliers for this attribute.

**Resident Status**

Resident status was coded with a value of 1 for residents and 0 for commuters. The mean value is 0.51 which in this case implies that 51% of the focus group are resident students, supported by the median value of 1. The standard deviation and variance are 0.5 and 0.25 respectively, which are low to be expected because the only values are 0 and 1, which explains the range of 1. The skewness is -0.041, very low suggesting a tight data set. Expectedly there are no outliers for this attribute and the quartiles are not worth mentioning.

**Extracurricular**

Based on the statistical analysis of whether students use the gym or play a sport, 45.9% said they do, given the mean of 0.459 and median of 0. This measure was also coded with a 1 for yes and 0 for no, so the standard deviation of 0.495, variance of 0.248 and range of 1 are expected. The skewness is 0.164, very low suggesting a tight data set. Not surprisingly there are no outliers for this attribute and the quartiles are not worth mentioning.

**Showering Location**

Based on the statistical analysis of whether students shower anywhere other than their place of residence, only 8.2% said they do, given the mean of 0.082 and median of 0. This measure was also coded with a 1 for yes and 0 for no, the standard deviation and variance are 0.274 and 0.075 respectively, so low due to how few values are not 0, also explaining the relatively high skewness of 3.056. The quartiles are all 0, supporting how the only values above 0 are in the top 25% of the data, therefore the outliers for this data were the cases where students replied yes.

**Average Showers per Week**

Based on the statistical analysis of the students average showers per week, the average amount of showers that students at UMass Dartmouth take is 6.612 given the mean. With a median value of 7 that indicates the data points are centered around that value and with a standard deviation of only 2.735 it is fair to say that the average shower time is fairly close to the median value. There were two outliers of 15 and 17 showers per week, which explains why there was a high range of 16. The two outliers are also a cause of the slight positive skewness of 1.034 as the data is centered around the median of 7 so there is an ever so slight trailing tail towards the right. The quartiles in order are 5, 7, and 7. This data shows that 75% of students surveyed take 7 showers or fewer a week and only the bottom 25% take 5 or less. That leaves the top 25% taking more than 7 showers per week.

**Average Shower Time**

Based on the statistical analysis of the average time spent showering of UMass Dartmouth students, the average shower took 16.031 minutes. The data centers around the median of 15 but with a high standard deviation of 9.238, there can be a point made that the data points are actually not too closely centered around the mean. Although what may be contributing to the high standard deviation is the large range of 58 caused by two outliers of 45 and 60 minutes. Once again there is a positive skewness of 1.739 which is most likely caused by those two outliers which have caused a slight trailing tail towards the right with most of the data points on the left. The first quartile for this attribute is 10, expressing that the bottom 25% of the population take a shower of 10 minutes or less, the second quartile is 15, therefore 50% take a shower of 15 minutes or less, and the third quartile is 20, representing how only the top 25% of students surveyed take showers of 20 minutes or more.

**Longer Showers**

One question that was asked was if UMass Dartmouth Students sometimes take longer showers than what they normally do. The responses were coded such that 0 represented that they do not take longer showers than normal and 1 represented that they do. 84.7% of the people surveyed said they do sometimes take longer showers than normal as indicated by the mean of 0.847. The median value supports this as most of the data points were centered around 1. With an expected low standard deviation of 0.36 and a variance of 0.13 as the only possible responses of 0 and 1 this explains the range of 1. The skewness of -1.927 reflects how most of the focus group answered that they did occasionally take longer showers. The quartiles are all 1, showing how only in the bottom 25% of responses did people answer that they never take longer than average showers. Expectedly, there are no outliers for this data

**Willingness To Reduce Shower Time**

Another broad question we asked was whether students believed they could reduce the time of each of their showers by one minute. This question was also coded with 0 as a negative response and 1 as a positive response. The mean calculated is 0.827, representing that 82.7% of the focus group believes they can shorten the length of their average shower. The median for this attribute is 1, supporting how the data is centered around a positive response to the question. Standard deviation and variance are 0.379 and 0.143, low as a result of the range being only 1. The skewness of -1.725 is consistent with the fact that a majority of the focus group answered that they believe they can reduce the time they spend in the shower by at least one minute. The quartiles are all 1, showing how only in the bottom 25% of responses did people answer negatively, and there are no outliers.

**Correlation**

|  |  |
| --- | --- |
| Age vs Average Shower Time: | -0.230966 |
| School Year vs Average Shower Time: | -0.258084 |
| Resident Status vs Average Shower Time: | 0.045231 |
| Extracurricular vs Average Shower Time: | 0.081178 |
| Showering Location vs Average Shower Time: | -0.065537 |
| Showers Per Week vs Average Shower Time: | -0.011645 |
| Occasionally Longer Showers vs Average Shower Time: | 0.213098 |

Based on the correlations calculated some relationships between the general attributes and shower times in minutes can be made. Age and school year have relatively high magnitude, negative correlations with average shower time, suggesting that older students tend to take shorter showers. When looking at whether students acknowledge occasionally longer showers, the correlation to average shower time is relatively high and positive, suggesting that those students tend to also take longer showers over all. Less significant outcomes show positive correlations with resident students and those who exercise or play a sport. Lastly, number of showers per week and whether students shower anywhere other than their residence seem to have negative correlations with average shower time.