**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 following document will detail the results of our initial focus group, mainly the changes that were made to and are going to be made to the survey itself. The changes include the content of the survey such as additional questions and changes to the already present questions. The other types of changes would include the overall distribution of the survey and what was done in that area. Along with that, it will describe a set of entries from the focus group and why they were omitted from the analysis portion as well as what their occurrence means to the project as a whole including examples of just how skewed the data can get.

**Focus Group Findings**

****

From Phase 1 to Phase 2 of the project, a handful of changes were done. Content-wise the survey did not change too drastically. One of the changes to the content was the addition of an ending question that asked the participants whether or not they could reduce their showering habits by at least a minute with just a simple *Yes* or *No* option. Barring that, not many other changes were made besides rewording some questions to flow better. The biggest change unto the survey was how it was to be distributed. Previously the survey was distributed digitally as a word doc or had their answers jotted down by simply showing them the questions, this method proved to be ineffective with the speed at which the results were being gathered. In order to mitigate the time wasted gathering results, the survey was transferred to a *Google Forms* version, which is attached via a link [here](https://drive.google.com/open?id=1NQ-mDy52KHEAyGQ8ujgyVxXa6MSVnmnqtMR5Yf6OoFc). Accompanying this new Forms version was a QR code by which to increase the effectiveness even more so (as seen above). We could walk over to potential participants and have them scan the code, if they were willing, and leave them to their own devices to fill out the survey.

The design of the Forms version had also presented some problems, both major and minor. Many of the questions had fill-ins rather than multiple-choice options. For example, with a question such as **what is your major** some of the results produced were acronyms and if they could not be properly decoded then they would just have to be assumed what they mean, creating a potential error in the final results. Similarly with that would be potential spelling errors of major titles, however this is only a minor grievance as spellcheck can sort that out fairly quick, but nonetheless it takes up valuable seconds that can be used on other processes. A much more major issue comes in when the participant of the survey does not take it seriously and in the fill-ins adds only joke answers that if used in the final results can skew the outcomes drastically. A proposed change to many of the aforementioned problems would be changing all of the fill-ins to drop-down menus, eliminating the need for the participant to spell, the need for us to decode any acronyms and to avoid any "joke" entries. Removing the fill-ins from some questions is also a possible change. This change would be used on a question like **what is their gender** and this comes about due to the "joke" entries as well. Our selections for gender in the current version are *Male*, *Female*, *Prefer Not to Say* and *Other* with *Other* allowing for the participant to fill-in what that other is. This fill-in also caused a variety of ridiculous answers like some identifying as inanimate objects.

**Omissions**

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Entry 1** | **Entry 2** |
| Gender | Lizard Fluid | Mail |
| Age | 9000 | 2.14756 |
| School Year | Graduate Student | Senior |
| Resident/Commuter | Commuter | Resident |
| Gym/Sports | Yes | Yes |
| Shower other places | Yes | Yes |
| Avg. Showers per week | 9000 | 1.123456789 |
| Avg. Shower time | 9000 | 1.123456789 |
| Occasionally longer | Yes | Yes |
| Can reduce | 9000 | N/A |

With the original focus group for this phase, not many of the entries were omitted from the calculations, however, the 2 seen above were. As mentioned in the **Focus Group Findings** section, some entries were exclusively made by participants as joke answers. These are the answers by which that statement was referring to. They waste our time in terms of analysis because they have to be removed from the final calculations. Luckily, the answers were back-to-back, so searching for them and removing them from the findings sheet was an easy task. Another good thing about these 2 entries is that they are the only ones of that variety in terms of ridiculousness. If the entries were not removed and were to be accidentally included in the statistical measures, then the data would be thrown off by a very large margin. For example, let's take the mean for various attributes and how they differ when the entries are left omitted compared to when they are left in. When they are omitted for an attribute like **Age** the mean sits at 21.112, but left in the mean increases by a large degree to 110.711476. This is an incredible 424.401% increase. Another attribute that would be drastically affected by leaving in those entries would be **Average Showers per Week**. When omitted, the mean sits at 6.612, but when added back in it jumps to 96.491235. This is a staggering 1359.34% increase. Those 2 increases alone present the danger of keeping those entries in.