**Statistical Design Consulting**

**SEMESTER REPORT**

**Fall 2024**

**Client:** Akhere Olenloa  **File Number:** 24-004

**Department:** Agricultural Engineering **Major Prof:** Klein Ileleji

**Consultant**: Sumeeth Guda **Initial Meeting Date:** 01/30/24

**Meeting Attendees:** Akhere Olenloa, Sumeeth Guda, Dr. Klein Ileleji, Dr. Bruce Craig

**Statement of Problem:** To investigate the challenges that grain farmers and grain elevators face regarding adopting grain monitoring technology. Additionally, to develop a predictive model for the adoption of the technology.

**Goal of This Project:** Ph.D Dissertation, Journal Article

**Background:**

The purpose of the client’s study is to understand the state of adoption of grain storage monitoring technologies. Evidence suggests that there is a very low adoption rate despite advancements in technology. The problem the client wants to address is why only 10% of on-farm storage owned by farmers use a form of monitoring technology, and 30% of off-farm storage owned by grain elevators use a form of monitoring technology.

The client created a survey to send to both farmers and grain elevators to determine what factors ultimately influenced the farmers or grain elevators to adopt or not adopt the grain monitoring technology. Within their survey they had 5 key areas:

1. The first section is collecting demographic information about the participants. Collecting the education level, gender, and region the participants are located in.
2. The second section is about the characteristics of the grain storage facilities that the farmers or grain elevators have. It asks about their storage capacity, the type of grain collected, storage period, business period, and if there are already grain monitoring technologies in place.
3. The third section is about the technological features and drivers for adoption of the grain monitoring technology, assuming the participant already utilizes grain monitoring technology. Specifically what company their technology is from, what factors are being monitored (Humidity, CO2, Spoilage, Insects, Temperature, etc.), the frequency of the monitoring, and the use of the factor within the grain management.
4. The fourth section collects data regarding the participants’ perceptions and benefits of adopting grain storage management technologies.
5. The fifth section is asking about the challenges and constraints with respect to the adoption or use of grain storage technology.

From these 5 survey areas, the client wanted to use the data collected to answer the following research questions:

1. Is there an association of the demographic information (ex. age and gender), grain storage duration, grain storage capacity and location of grain storage on the adoption of stored grain monitoring technologies among grain elevators and farmers?
2. Do grain elevators and farmers’ experiences with managing stored grain influence their adoption of stored grain monitoring technologies?
3. How do grain elevators and farmers’ perceptions of stored grain monitoring technologies influence their adoption of stored grain monitoring technologies?

The ultimate end goal of this survey was to create a regression model to predict the adoption of grain monitoring technologies and to determine what factors are significant in the adoption.

**Progress During Current Semester:**

For the most part this semester, the client was focusing on another project for his research. But he worked with the consultant periodically to get help for this project. The most important update on his end was that he was finished collecting the data for his survey and was now ready to focus exclusively on the analysis of his collected data instead of toy examples.

The client and consultant worked this semester to refine the analysis approaches as discussed in the spring semester IM. But as seen in the summer semester, one of the big issues with Akhere’s data is that he only has 37 responses from the survey and over 100 predictor variables. The model ended up having a lot of lack of fit issues since the model space has a high dimension and the model itself is under saturated.

Sumeeth asked Dr. Craig for his insight into what Akhere can do for his analysis, Dr. Craig’s suggestions were the following:

Dr. Craig’s suggestions:

He acknowledges that you are in a tricky position with your data given that you only have 37 cases, hence you can only consider models involving subsets of predictors (combination of predictors, possibly with interaction terms. But not using all ~100ish predictor). Dr. Craig felt that is If the responses are either ordinal or binary, then the problem becomes more complicated since you must deal with separation of points.

Some of the possible solutions he recommended were:

1. Incorporating a penalty (like Firth’s) is one way around separation of points. (This will also prevent overfitting within your model)
2. Not all association questions require all the predictors in the model.  For the association questions, consider which predictors are your primary predictors and make the other variables covariates.  Consider breaking things up in this way to answer your questions?  Additionally consider looking at associations with and without the covariates.
3. Dr. Craig did mention that dimension reduction could also be used.  His recommendation was to apply PCA (Principal Component Analysis) to the covariates and use the first couple of dimensions as the predictors.

Overall, the consultant and client felt that having a follow up meeting for the analysis phase could be beneficial. However, since Dr. Ileleji was on sabbatical, a follow-up meeting would have to happen in the spring semester.

**Current Status: Continuing, Follow-up needed**