**YouTube Speech**

**Module 8 Capstone Project Final Report Option 1 United States Organization**

“Data has shape, and shape has meaning.”

**GREETING/OPENING**

Hello, Dr. Mills and my esteemed audience.

My name is David Jurist. I am pleased to have this opportunity to visit with you to review my Capstone Project for MIS581 at CSU Global.

In this brief presentation I will address the rubrics of the course to show how this project fulfills the course requirements, and I will address the full spectrum of elements that support the USDA’s public stakeholders with my vision that matches your own.

My project reviews the main components of the USDA organization, research hypotheses, complex literature review, data sets, analysis, findings, conclusions, and recommendations for valuable future work.

**INTRODUCTION**

You will notice that my scholarly references are all accounted for at the end of the PowerPoint slides and in the Capstone writeup.

The theme of the writeup focused more on the overall research, data, and analytics. The theme of the PowerPoint slides focused more on the business and organizational perspective. And this presentation focuses more on the highlights of the work as a finite project.

What we wanted to accomplish in the project is to open the process of research and full data analytics lifecycle for a model and simulation of the global food chain. A very ambitious undertaking that will take much time and effort to accomplish.

Data analytics project, yes, but research project, with much discovery, adoption of tools, alignment with strategy of the USDA and the users of its data, and a clearer understanding of what all this work will require.

**HYPOTHESIS**

My research hypothesis is that it is feasible to begin modeling the global food chain. It is begun with a consolidated data set for all food crop production correlated to the price of energy required to produce the food. The hypothesis is that the volume of global production of food grain crops varies inversely with the average annual price of West Texas Intermediate crude oil, which represents the cost of energy applied in the agriculture industry.

The null hypothesis H0 states, when N = 32 and r is the correlation coefficient, r = -1.00

The alternative hypothesis HA states, r != -1.00

Another hypothesis is that the world population grows in response to the increasing production of food, not vice versa. This hypothesis will require additional hypothesis testing after this project with the food production and world population statistics already gathered.

**LITERATURE REVIEW**

In the literature review I discovered that ample literature is available, but not necessarily about the wholistic, global perspective, since food production is predominantly a local activity, and statistics collected nationally.

Two treatises that most influenced my thinking were the volume by Gary Kleppel (Kleppel, G. (2014). *The emergent agriculture: farming, sustainability and the return of the local economy*. British Columbia, Canada. New Society Publishers.) and the accounting of the NABE survey that discussed how institutions are combining government data (including USDA data) with their own Big Data with data analytics for business value formation. Such data are used . . . for purposes that include production and investment decisions, marketing and inventory management, and long-range strategic planning” (Hughes-Cromwick, & Coronado, 2019, p. 145).

“A new kind of farming is emerging in America” (p. XIII), to improve our “prospects for revolutionary change in our food system” for sustainability and nutrition.

**DATA**

Since the data is USDA data, it predominantly describes United States agriculture. Data from other nations is not readily available, except for the export/import data prepared by the USDA.

The data sets selected were limited to the food-grain crops only, WTI crude oil historical prices, and the world population

All the data was time series data.

Slogan: “Data has shape, and shape has meaning.”

**ANALYSIS**

Although I calculated them the descriptive analytics would not be entirely meaningful in this project with time series data (such as the mean, mode, median, SD, variance) for 32 years of production and energy prices. Regression will be much more significant for predictive estimates of production and prices.

Instead I focused on the correlation coefficient.

For examining, cleansing, formatting, converting, and preparing the data I used the text editor in Weka, WordPad and Excel. Analytics were done in R CLI, Weka, and Tableau.

**FINDINGS**

Test of the null hypothesis failed, as the calculation resulted in r = -0.4729

The initial null hypothesis was disproven by the data not demonstrating close inverse proportion.

The writeup disclosed detailed observations about the data and analytics, but the highlights are:

The test disproved the null hypothesis.

1. More precise data about energy resources consumed solely by the agriculture industry and for specific crops will reveal better test results in the future.
2. Much data is available for the United States, but little is available for other countries. Little has been analyzed and published by the USDA except with economic reports, but much data is available.
3. Since food production is specifically a geographical activity, territorial and national mappings of the data will make for more revealing visualizations.
4. Security and privacy are not prima facie a concern.
5. The United States is the largest food producer in the world.
6. Blockchain technology (Bedord, 2019) with IBM will enable new applications of technology for the decisions that will help feed the poor and enable the producers who can feed the poor.

**CONCLUSIONS**

Continue with the project, add more analytical functions and algorithms, and test more hypotheses.

This project only scratched the surface.

I keep a strong purpose and motivation for discovering causal relationships and interdependencies in the model.

An interactive, predictive-prescriptive model of global food is feasible, but will require long-term construction of the fully integrated model with its contributory sub-models of energy, weather, natural resources, fertilizer and pesticide.

“The value of government data is difficult to measure, but it is clearly a substantial strategic asset for the US business sector” (Hughes-Cromwick, & Coronado, 2019, p. 145).

Ultimately the intention of analytics is to arrive at insights from descriptive, predictive and prescriptive analytics that extract that intrinsic value of United States government data for actionable business, financial, and agricultural decisions.

**FUTURE WORK**

Water, weather, energy, fertilizer, pesticide

Flow cradle to grave, distribution, economic and investment decisions, Export/Import.

Make it online, interactive, colorful, geographical, predictive-prescriptive, real-time, current.

Experiment with AI for innovative integrated solutions and perspectives

Build a data warehouse in the cloud

Expeditiously use cloud resources with innovative online user applications

Assist prospective clientele with sophisticated planning services and online, self-service tools

Continue to find and gather Big Data from all sources and all types: Structured, unstructured, created, provoked, compiled, text, video, audio, image, Internet of Things

Comprehensively assess predictive, descriptive and prescriptive algorithms linked to USDA strategies

Explore social media Big Data that would contribute value to these initiatives

Advance macro-economic and financial models with source data for investment, financial, credit, and operational decisions.

**CLOSING/GOODBYE**

As we prepare to close this presentation, allow me to say that this was a very rewarding and educational project for me, and I do hope that it has enlightened you as well.

I discovered this is a research and business value-formation endeavor, not just analysis, statistics and programming.

I reviewed the research hypothesis, the data, the analytics, the findings, the conclusions, and recommendations for continuing research and analytics.

I reaffirmed the value of the full data analytics lifecycle.

Remember that data has shape, and shape has meaning.

Until the next time, enjoy abundant results in all you do.

Thank you very much for your support, direction, professional consideration and attention, and goodbye.