**Step 1: Big Data Semester Project – Your Pitch! Due next Friday by 4 PM**

This will be graded as a final, akin to last semester’s project. Here’s your chance to really show off what you can do! Embrace the opportunity to explore a topic of interest, collaborate, and do some college-level research. 😊 ALSO, test your ability to manage your time, stay organized, and be assertive about issues that come up.

Collaboration

Describe ONE of the following.

* Work with a partner. Mrs. Frazier will grade you in class based on your communication, effective use of time, and problem solving/discussion about conclusions together – tell Mrs. Frazier in a few sentences (that you and your partner can write together) why you know your partner will be a good resource for you (be specific). If you work with a partner in class, you need to turn in TWO pitches, with both names on the top; working with a partner does not make you have less work than working as an individual with an expert. This means you will each do Big Text analysis, for instance, on different (but related) topics that can help provide insight of their own. Your code can use results of another student’s code/work together, but you’ll want to be sure to have plenty to say about your methods that you made yourself. Likewise, you can choose to write your conclusions about your Big Text, Statistics, Meme, and ArcGIS together; each of these don’t need individual paragraphs unless you and your partner decide that is best, but you would each contribute your own graphics of different things for each of those.
* Work with an expert. This expert will work with you weekly, providing written feedback (and a photo of you with your expert) about your work. Your expert needs to be able to help you understand what you are seeing in your data, so this person should know a lot about *a, x, and y,*  below. Describe who your expert is and what insight they can provide that you don’t have. Specify when you will meet with your expert, once per week, for at least 15 minutes, and be sure your expert knows they will have to write you SPECIFIC feedback (list form is fine) and provide an in-person photo with you at each meeting.

Need

**Create a strong need statement** in which:

* Your audience (*a*) is clear and has a clear drive/emotive angle.
* There is fantastic insight. Remember, obvious is not insightful.
* Behavior or what is needed is clearly rationalized.
* There are at least two clear variables you are suggesting that are related in some way (*x, y*).

*Think back to one of our earliest Need Statements … about teens and healthy eating. You should be able to identify all features above in that.*

 ***Be sure you define a, x, and y clearly. You use them more below.***

Prior Research

**Write a list including full citations** (include footnote and bibliography) summarizing what is already known about the relationship between:

* *x* and *a*
* *y* and *a*
* *x* and *y*

*This is a check for obviousness. If the world already KNOWS about x having a relationship with y, then you need to back up and revise your Need Statement. If the world only suspects the relationship between x and y, you can keep going – your research CAN explore something other people think.* ***You should not need to spend more than an hour finding research and summarizing it in this section. Some of this might help you fill in ideas in the below table, as well.***

Data Sources

Each data set needs to be BIG and the right *kind* of data (for instance, Big Text Analysis cannot be done on numbers). Complete this table, being sure you explore all variables and combinations of variables at least once:

|  |  |  |  |
| --- | --- | --- | --- |
| Work to do | *x*, *y*, *a*? (list all that apply) | Summary of what you are exploring (a focus question or hypothesis, for instance) | Data Source – description and full bibliography |
| Big Text Analysis (word clouds & Chi-squared Test) |  |  | 1)2) |
| Meme Analysis (>1 OSoMe tool) |  | *Include the term time plot or network map.* | *Indicate > 2 memes you’ll try (no need for a bibliography)* |
| Geographic Information Systems Mapping (ArcGIS) |  |  | 1) 2)  |
| In-sheet Statistic calculation (Visual Basic or Google Apps Script) |  | *Mention statistic!* | *(> 1)* |
| High-Level Statistical Graph(recommend linear regression, x vs. y) 🡨 this can come from a research study already published! |  |  | *(> 1)* |

Coding with a Collection and Input to provide an Output

This is the CREATE portion of the final. Be sure your code meets an objective that others would be interested in, and is original. You can use any computer language you would like.

Describe in a sentence what each method will do. You need 2 methods that you write yourself,\* each method must do something interesting and useful related to the bigger picture,\*\* and one method must require another method to run.\*\*\* At least one of your methods should be working with your abstraction you specified above.

\* No credit if you use a pre-existing method such as *print*, and add in parameters. “Write yourself” means you write the header and body of your method yourself. You must work with input from either: user, device, data stream, or file, and produce some kind of clear response. See <https://apcentral.collegeboard.org/pdf/ap-csp-student-task-directions.pdf> pages 4-10.

\*\* Initialization, declaration, running your other two methods, and printing are not, on their own, interesting. The body of ONE method should include some processing that uses ALL of the following: sequence, selection, iteration. You’ll demonstrate the selection during execution by providing your method with different parameter(s) and showing different responses/return values.

\*\* ANOTHER method needs to use a collection of some kind in a way that clearly shows off how that collection helps to manage complexity and is far more efficient than NOT using a method (you can describe this theoretically).

\*\*\* This often means that the second method relies on the result from the first method, but there are other ways to make this work.

1. Method 1 is called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and does this:
2. Method 2 is called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and does this:

Timeline

Describe how you will divide the work, establishing a timeline by which to turn in work. Be thoughtful! Some work has to finish before other work begins! You can do this as a list or on a calendar.

**Step 2: Work time in class. Mrs. Frazier will offer peer edit time, rubrics, and guidelines about how far you should be. 😊**

**Step 3: Final deliverables to turn in on May 19, NO EXCEPTIONS.**

Collaboration: Mrs. Frazier will grade OR you’ll turn a photo and written feedback by each Friday. On the final day of work, please write a paragraph reflecting on collaboration in this project, including (a) what you gained from your partner (and what you lost, if that came up), (b) a list of tips you learned from working with your partner, and (c) (for in-class partnerships) any amendments to how your split up work or comments about how Mrs. Frazier should grade you, if grading is not expected to be equitable.

Need & Prior Research: nothing to change here if all was fine with the pitch; just resubmit the full pitch again.

Coding to Process an Abstraction: (a) Create evidence of checking for Completeness (this should be made during the design portion of your coding). (b) Check/describe all 6 qualities of code - be specific, providing examples as necessary. (c) Follow ALL directions for College Board’s Create. There is an excellent peer review tool on the Weebly site for class, which you should use when editing your work.

Analyses and Calculations: These will work like our smaller projects in class. Cite sources of data and provide output of analysis (often a graphic) and a paragraph describing what can be concluded. Use specific evidence in your paragraphs and use appropriate terminology like “co-occurrence.” Use a null hypothesis, hypothesis, or focus question as a subtitle preceding each analysis. For instance: *Big Text Analysis: Do fleas worry about their next meal?*

Big Text Analysis

Meme Analysis

ArcGIS Analysis

In-Sheet Statistical Calculation: In addition to the above general directions for analyses and calculations, include screen shots of before AND after showing that the sheet is dynamic. Also click on cells that are programmed to show their **syntax** – you can show this in screen shots or label up a graphic with arrows and note what cells did what. Include the name of the statistic you calculated in the conclusion paragraph.

High Level Statistical Graph

Summary: Restate your Need Statement, then write a paragraph summarizing what you learned, citing specific evidence you found most compelling. Suggest what future research could be done, either to clarify things or expand your topic of research to another level of insight (audience and/or variables can change).