**ANLY503 Exam Take-Home Portion (100 points)**

ASSIGNED Friday by 6pm ET

**UPDATED:**

**DUE Thursday Dec 7 at 11:59 pm ET (final deadline with no penalty is 12/9 by 12 noon ET)**

**Grace Period: I will accept this portion through Sat. 12/9 by 12 noon ET. This grace period is optional. Technically, I cannot make this due past Thurs - so I am not changing the due date, I am just allowing late submissions with no penalty through 12/9 at 12 noon ET.**

**NOTE: Dear Class - I am going to allow more time for this given several reasons. Please note the new due date and the fact that this cannot be late.**

**IMPORTANT NOTE: I know that this dataset is on Kaggle and so there are many visualizations that already exist for it. Do not “borrow” them - make your own. Please do not undermine my exam or you will waste your time and my time ;) This is an exam to see that you can make visualizations (not find them).**

**Directions:** This take-home exam will check to see if you worked on your own Portfolio, coded your own visualizations, and learned how to create excellent visualizations that can describe data. This exam is \*intended\* to allow a small and finite amount of time to see what you can do given a dataset and a time constraint. You can use ANYTHING that you have created, ANYTHING that I have shared with you, and/or that you have made in your Portfolio or during class. This is also open Internet as well.

**You may NOT discuss this with any humans.**

**What you need to do:** Storytelling with Visualizations

Use [THIS DATASET](https://drive.google.com/file/d/15PEJtDR3nzrczWPUZZpIWSrHlDOkft4F/view?usp=sharing)

Inside the dataset you will find **a link to a webpage link (pdf) that describes the variables** in this dataset. It is **not** necessary to use all variables. You will decide which ones to use and you will decide if and how you may or may not need to clean things up.

**Overall Goal 1:** Using EDA for exploration with static graphs.

**Overall Goal 2**: Create a web page (online) with only interactive visualizations that **tell a story** about this dataset.

1) Use EDA to explore the data. Create five (5) EDA non-interactive visualizations (in any language) that you think will help \*you\* to better understand what this data is about. Create a **Word doc.** (only) that includes these EDA visualizations (paste them all into the Word Doc). Create 10-15 sentences (in total - not per vis) that concisely describe what all of the EDA graphs have helped you to learn about the data. (Again - not 10-15 sentences for each graph, but rather 10-15 sentences in total that summarize what you have learned in your exploration). **(20 points) You will submit this as .doc or .docx to BlackBoard. Call it EDAYourName.doc(x)**

**Making the Website Visualization Story: (80 points in total)**

1) Determine what story \*you\* think this data is telling. Write 3 - 5 sentences that clearly state what you want to show and tell - what your story is. Place these sentences under your title and at the top of your webpage. **You may use ANY option** to place this story online as long as you meet the requirements. (10 points)

**See below for the rest of the requirements.**

All of the rest is worth up to 70 points

This assignment is your chance to show that you can choose and create smart, interesting, and appropriate interactive visualizations that can **tell a story about data**.

\*You\* will determine the types of the visualizations and the details of each visualization - except for the following minimum requirements:

**Minimum Requirements:**

- At least one network vis (interactive) Yes - you must use this data to make the network vis and so you may have to manage the data and choose variables so that you can make a network that makes sense. (10 points)

- At least one geo vis (map) (interactive and must be done in Leaflet) (10 points)

- A least one geo-temporal vis (maps over time using a subplot) – this subplot of maps over time will count as one (1) of the required graphs, even though it will be a set of graphs in a subplot. (10 points)

- At least one temporal vis (5 points)

- **At least 10 visualizations in total**  - all interactive - including the above. So to be clear - the above visualizations are four of the ten required graphs.

- You can use Tableau for up to two (2) of the graphs. If you use Tableau for the geo-temporal set of subplots - that will only count as one of the 2. You cannot use Tableau for more than two of the graphs.

NOTEs: **Avoid bar graphs and pie graphs.** If you are not sure that something counts as a bar or pie - just do not use it to be on the safe side.

\*\*Each graph/vis should be very helpful and useful in representing key element(s) of the data.

All visualizations should be fully and properly labeled.

For each Vis, you may include 1-3 sentences (no more) that describe the graph.

**For each vis - include a link to the code and the data.** If you created an updated or reorganized version of the dataset (such as you will have to do for the network graph) include that instead of the dataset. In other words, include the data that your your code uses to make the graph.

**Other than the noted requirements above, you can use any other packages that you choose, including, but not limited to Python3, R, NetworkD3, Plotly, Bokeh, etc. etc.**

**Note that up to 20 points are set aside for the look and feel, clarity, flow, and excellence of your Visual Story. This is subjective so think about how to make me understand your story.**

**Submit the URL to this Story to BlackBoard. Place the URL right in the comments area AND place it in a Word doc - call the Doc TakeHomeANLY503ExamYourName.doc(x).**

**How to Submit**

Item 1: EDA Word Doc

Item 2: Word doc with URL

Item 3: This document that you are required to read and place your name on below.

Place all three items in **TakeHomeExamYourName.zip** and submit to BlackBoard

**NOTES:**

To review, your goal here is to use (and clean as needed) some or all (your choice) of the attributes in the [dataset](https://drive.google.com/file/d/15PEJtDR3nzrczWPUZZpIWSrHlDOkft4F/view?usp=sharing) and to tell a clear and concise story about this data. You will decide what to use and how to use it. You will determine the visualizations that do the best job in describing the data and telling your story. You will determine everything about this assignment other than the basic requirements noted above.

**Grading:**

This Take-Home Exam is intended to see what you can do with a given dataset and limited time. I will be graded using a variety of evaluation measures. Even though there are points associated with each requirement above, you can also lose points for not meeting any of the following:

These will include, but will not be limited to things like how creative your visualizations are (don’t be boring), how well they describe the data, your use of colors, shapes, styles, and positioning, your use of labeling, axes, titles, and subplotting, and your general ability to use visualizations to tell a story. I will be looking at the “flow” of your webpage (the URL you submit). I am expecting it to be self-explanatory, easy to understand and follow, and for it to tell a story as a I scroll. **This is a creative assignment. No two submissions will be similar. You cannot ask any questions of me or anyone else. \*You\* make the decisions – that is part of the exam.**

BE CAREFUL – visualizations and webpages are like fingerprints – yours should not look like anyone else’s. DO NOT share or discuss this so as to avoid also sharing the grade (at 50% each). This is NOT a shared activity or group activity - this is an exam. DO NOT share or discuss this exam with anyone.

**Type your name here to agree that you completed all parts of this exam without any other human assistance.**

Name Yuan-Yao Chang - Date 12/08/2017