DAV 5400 Fall 2019 Final Project

Deliverables Schedule

Deliverable	*Date*	*Points*
Proposal	1st Draft due No later than **Friday Oct 27;** Final Draft due **Wed Nov 20**	30
Final Project	Wednesday Dec 18	170
Final Project Presentation	Before or during Final Class on Monday Dec 16	50

Policy on Collaboration

You may work in a team of up to three people. Each project team member is responsible for understanding and being able to explain *all* of the submitted project code. Remember that you can take work that you find elsewhere as a base to build on, but you need to acknowledge the source, so that I base your grade on what you contributed, not on what you started with!

Proposal Guidelines

Your proposal should adhere to the outline and guidance provided in the separately provided **ProposalGuidelines** Jupyter Notebook.

Approval Meeting

Once you've turned in your proposal, I want to schedule a brief phone meeting with each team to discuss the proposed project. For team projects, I also want you to articulate the roles and responsibilities of each team member.

Final Project Checklist

To receive full credit, you'll need to deliver on all of the items in the checklist below. **Please read carefully through this checklist before you make your project proposal**. You are (within these checklist constraints) strongly urged to limit scope and make the necessary simplifying assumptions so that you can deliver your work on time!

- ☐ Proposal describes your motivation for performing this analysis.
- ☐ Proposal describes from where you plan to source your data.
- ☐ Your project has a recognizable and reproducible "data analytics workflow." [Example: First the data is acquired and explored, then necessary transformations and clean-up are performed, then the analysis and presentation work is performed]
- ☐ Project includes data from at least ***two*** different *types* of data sources (e.g., two or more of these: (1) relational or csv, (2) scraped web page, (3) web API.
- ☐ Project includes statistical analysis and graphics that describe and/or validate your data (e.g., EDA).
- ☐ Project includes at least one data transformation operation. [Examples: transforming from wide to long; converting columns to date format]
- ☐ Project includes at least one grouping or aggregation.
- ☐ Project includes at least one graphic that supports your conclusion(s).
- ☐ Project includes at least one statistical analysis that supports your conclusion(s).
- ☐ Project includes at least one feature that we did not cover in class! There are many examples: "I created a decision tree; I ranked the results; I created my presentation slides directly from my Jupyter Notebook; I figured out to use OAuth 2.0…", etc.
- ☐ Presentation. Was the presentation delivered in the allotted time (3 to 5 minutes)?
- ☐ Presentation. Did you show (at least) one challenge you encountered in code and/or data, and what you did when you encountered that challenge? If you didn't encounter any challenges, your assignment was clearly too easy for you!
- ☐ Presentation. Did the audience come away with a clear understanding of your motivation for undertaking the project?
- ☐ Presentation. Did the audience come away with a clear understanding of at least one insight you gained or conclusion you reached or hypothesis you "confirmed" (rejected or failed to reject…)?
- ☐ Code and data. Have you delivered the submitted code and data where it is reproducible and self-contained—preferably in a Jupyter Notebook on GitHub? Am I able to fully reproduce your results with what you've delivered? You won't receive full credit if your code references data on your local machine!
- ☐ Code and data. Does all of the delivered code run without errors?
- ☐ Deadline management. Were your draft project proposal, project, and presentation delivered on time? Please turn in your work on time! You are of course welcome to deliver ahead of schedule!