DATA 606

Goals for the Course

This Semester:

- Understand what your prospectus and oral defense will entail.
- Write sections of your own prospectus.
- Provide feedback and critique on your colleagues drafts.
- Learn how to use tools such as ChatGPT effectively and legally.

Cadence of the Course

We will meet as a full group every **Tuesday** (12:30 - 1:50PM).

Most weeks, you will then additionally have an assignment that requires you to get feedback from your advisor. Some weeks, you may have a one-on-one with me as well, which we will schedule as may be useful/necessary.

Grading in the Course

We will meet as a full group every **Tuesday** (12:30 - 1:50PM).

Most weeks, you will also need to meet with your advisor to make sure you are on track.

Your grade will ultimately be based on two things:

1) Your introduction, literature review, and quantitative analysis drafts (50%)

2) The quality of feedback you provide to your peers.

Timeline

October 1: Introduction and Draft Literature Review Due

October 8: Intro / Literature Peer-Reviews Due

October 15: Revision to Intro/Lit review Due (Dan & your Advisor)

November 12: Quantitative Analysis Drafts Due

November 19: Quant. Peer-Reviews Due

November 26: Revision to Quant due (Dan & your Advisor)

December 3rd: Full draft (Intro, Lit review, Quantitative section) due to Dan

Important Notes

- I am available for office hours, but they will need to be scheduled in advance. Please reach out to me if you need a one-on-one.
- I expect you'll meet with your advisor at least once every two weeks (minimum) to discuss your prospectus, and you will submit a draft to them at the same time as you submit it to me. It is critical that they are behind your preliminary ideas (far more important than my opinions!!).
- I am primarily here to guide you on the _procedural_ elements of your prospectus. I won't be able to provide enough _substantive_ feedback on, for example, the specific algorithms your proposing - you will need to work closely with your advisor on that (I'll provide some, but my specialty area is slightly-outdated computer vision techniques!).

Pause for Q&A

What is a Prospectus?

One of three parts of your qualification exam.
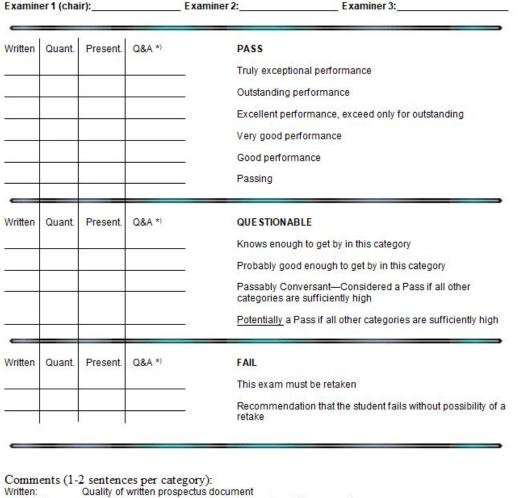
Part 1: A written exam (3 hours, pencils only). Late Feb/March.

Part 2: Your written prospectus (focus of this course). Due in April.

Part 3: Your oral defense (part of the focus next semester). Late April/May.

Who grades my Prospectus, and how?

Your prospectus is graded by three faculty members, none of whom are allowed to be your primary advisor. The grading scale is the same for all three sections (Written, Quantitative, and Presentation). Q&A is considered a part of your presentation, but broken out separately on our score sheets.

Examiner 1 (chair):_____   Examiner 2:_____   Examiner 3:_____

Written	Quant.	Present.	Q&A *)	
				PASS
_____				Truly exceptional performance
_____				Outstanding performance
_____				Excellent performance, exceed only for outstanding
_____				Very good performance
_____				Good performance
_____				Passing

Written	Quant.	Present.	Q&A *)	
				QUESTIONABLE
_____				Knows enough to get by in this category
_____				Probably good enough to get by in this category
_____				Passably Conversant—Considered a Pass if all other categories are sufficiently high
_____				Potentially a Pass if all other categories are sufficiently high

Written	Quant.	Present.	Q&A *)	
				FAIL
_____				This exam must be retaken
_____				Recommendation that the student fails without possibility of a retake

Comments (1-2 sentences per category):
Written: Quality of written prospectus document
Quantitative: Separate evaluation of the quantitative section of the prospectus
Presentation: Quality of the presentation (slides, presentation style)
Q&A: Level of the student's handling of the question & answer part following the presentation

What Makes Up a Prospectus?

Prospectus documents are generally around 50 pages long, and **must** contain a section explicitly dedicated to a quantitative analysis **you** have already conducted. This does not need to be novel (i.e., you can recreate a previous experiment), but it can be.

A normal Prospectus Outline

- Introduction
- Literature Review
  - Overarching Themes
  - Chapter 1
  - Chapter 2
  - Chapter 3
- Quantitative Analysis
  - Data
  - Methods
  - Results
  - Discussion and Conclusion
  - Supplemental
- Chapter Roadmaps
  - Dissertation Paper 1, 2 and 3 (you will have 3 different sections, just not duplicating for brevity)
    - Major Research Question & Intro
    - Proposed Data & Methods
    - Initial Results and/or Anticipated Challenges
- Timeline for Degree Completion
  - A realistic estimate for how long it will take you to complete any experiments you describe, and write up your results.
  - A Gaant Chart illustrating the tasks required.

This semester we're focusing on

- Introduction
- Literature Review
  - Overarching Themes
  - Chapter 1
  - Chapter 2
  - Chapter 3
- Quantitative Analysis
  - Data
  - Methods
  - Results
  - Discussion and Conclusion
  - Supplemental
- Chapter Roadmaps
  - Dissertation Paper 1, 2 and 3 (you will have 3 different sections, just not duplicating for brevity)
    - Major Research Question & Intro
    - Proposed Data & Methods
    - Initial Results and/or Anticipated Challenges
- Timeline for Degree Completion
  - A realistic estimate for how long it will take you to complete any experiments you describe, and write up your results.
  - A Gaant Chart illustrating the tasks required.

Prospectus need to be done using Latex Template

I recommend Overleaf, but you can use any tex compiler you prefer.

The template files are here (forked from a dead repo past students have used): https://github.com/D8A-SCIENCE/WM-AS-thesis-latex-template

Contributions back to this that will make future students lives easier are very much welcomed.

Focus for the coming week

BY NEXT TUESDAY, you should write a very early draft three-page introduction to your prospectus. Please reference the examples I have given you.

IN CLASS TUESDAY, each of you will present your overarching theme and your tentative three primary research questions. Please connect with your advisor to discuss these with them before you come to class. If you don't have all three nailed down yet, that is most certainly ok - be creative!

Your intro should have a <u>very</u> brief literature overview which contextualizes the problem(s) you are trying to solve.

Open/ go over structure of Warnke's introduction.

Chapter 1

Introduction

Instances of social unrest, often manifesting as riots or protests, wield significant influence on the communities, regions, and nations in which they unfold (Bencsik, 2018). The repercussions of such events are wide-ranging, ranging from geopolitical transformations (i.e., riots in Egypt in 2011 (Joya, 2011) and Hong Kong in 2019 (Purbrick, 2019)) to substantial economic losses (exemplified by the hundreds of millions of dollars incurred during the 2011 riots in the UK (Bencsik, 2018)). These events may result in human casualties, as evidenced by food riots in Africa in 2007-08 (Berazneva and Lee, 2013) and riots caused by garbage collection issues in Beirut in 2015 (El Warea et al., 2019). These events impact cities across the entire globe, with recent examples in Latin America (Eckstein, 2001), Asia (Purbrick, 2019), Africa (Joya, 2011; Berazneva and Lee, 2013), and Europe (Andronikidou and Kovras, 2012). Because of the importance of these events, scholars across multiple disciplines have sought to both predict and understand them, using a wide range of data sources and techniques (Pond and Lewis, 2019; Snow, Vliegenthart, and Corrigall-Brown, 2007; Davies et al., 2013). However, most of these approaches have relied on sources that may not be available or reliable in geographies of interest, such as news articles. Here, we explore the capability of satellite imagery to aid in the prediction of protest and riot events, explicitly seeking to understand the degree to which this globally-available source of information may be able to augment existing predictive methodologies. This approach exploits correlations between the human built environment - i.e., urban form (Fox and Bell, 2016) - and the likelihood of a protest or conflict event at a given geographic location.

One of the core innovations that enables us to estimate social events (such