EVANS SCHOOL
OF PUBLIC POLICY & GOVERNANCE

UNIVERSITY *of* WASHINGTON

PUBPOL 543: VISUAL ANALYTICS FOR POLICY AND MANAGEMENT

Prof. José Manuel MAGALLANES, PhD.

Winter 2020

E-mail: **magajm@uw.edu**
Office Hours: 11:30 – 13:00 (Thursdays)

Website: **https://canvas.uw.edu**
Class Hours: 8:30 – 11:20 (Thursdays)
Class Room: Social Work Building (SWS) 032

Course Description

This course gives students the tools to show insights to political or scientific communities, while presenting different strategies to avoid biased interpretations. Given the overwhelming computational toolbox for displaying information, the course follows a 'keep it simple' approach from the beginning, starting from foundational topics relating color, nature of data, and the brain; and takes students to build their own visualization tools. Emphasis is placed on complex data such as networks, geography and multivariate models. While the course uses R (and some Python), it makes no emphasis on programming and more on the building of templates to produce information.

There is no pre-requisite for this course.

Course Objectives

1. Carry-out visual analytics following a reproducibility/replicability approach.
2. Master basic data visualization techniques following a 'keep-it-simple' approach.
3. Become an effective user of templates in R and Python to produce graphics.
4. Be a reflective user of data science tools for informing on public data.

Course Book

There is no required book for the course. You can supplement some classes with:

> Magallanes Reyes, J. M. (2017). *Introduction to Data Science for Social and Policy Research: Collecting and Organizing Data with R and Python*. Cambridge University Press, 1 edition edition.

I might provide online material for reading before class.

Recommended Readings

- Python Language Documentation: **https://www.python.org/doc/**
- R Manuals: **https://cran.r-project.org/manuals.html**

Software installations required

Students have to install the following software in their computers:

- ANACONDA by Continuum Analytics (choose according to your Operating System):
  https://www.continuum.io/downloads
- R (choose according to your Operating System):
  https://cran.r-project.org/
- RStudio Desktop Personal License (choose according to your Operating System)
  https://www.rstudio.com/products/rstudio/download/
- LATEX:
  https://www.latex-project.org/get/.
- ZOTERO:
  https://www.zotero.org, and then download the desktop app.
- GITHUB:
  Get an account at https://github.com/, and then download the desktop app.
- GEPHI: It is not strictly needed, but worth trying it if time is available.
  Get it at https://gephi.org/

Evans School Community Conversation Norms

This course has adopted the Evans School Community Conversation Norms. Please be aware of these norms in interactions with the instructor and other students. At the Evans School, we value the richness of our differences and how they can greatly enhance our conversations and learning. As a professional school, we also have a responsibility to communicate with each other–inside and outside of the classroom–in a manner consistent with conduct in today's increasingly diverse places of work. We hold ourselves individually and collectively responsible for our communication by:

- Listening carefully and respectfully
- Sharing and teaching each other generously
- Clarifying the intent and impact of our comments
- Giving and receiving feedback in a "relationship –building" manner
- Working together to expand our knowledge by using high standards for evidence and analysis

Changes to the Syllabus

The professor reserves the right to make changes to the syllabus during the quarter. The professor will notify students immediately by email and in class if any changes are made.

Grading Policy

Grades consider two elements:

- **Group Exercises**: Students prepare a project to be done in groups.
- **Individual Exercises**: Each session has a set of exercises to be completed in class. Each Exercise submitted after class will reduce the grade by 5 / 100 points per day of delay.

These elements are graded like this:

- **30%** Individual Exercises.
- **20%** Group Exercise 1: present a problem of study with multiple variables on the same unit of analysis (countries, counties, and the like). Propose two indexes to plot.
- **20%** Group Exercise 2: present the plots for the indexes proposed (univariate and bivariate).
- **20%** Group Exercise 3: Present a report from your group: format can be in a web page or paper-like format.
- **10%** Group Exercise 4: Present a member evaluation of member participation.

Working in groups

This course requires working in groups. Consider that *Homeworks* and the *Data Project* depend on data that you will select according to your needs or interests. Thus, working in groups will combine different knowledge and will bring more insight to the group. You should have decided who will be in your group by the end of the **third** week of classes.

Course Schedule

Week 01, 01/06 - 01/10: Visualization Basics

- Components.
- Objects.
- Color.
- Reproducible environments
  - Creating a Github repo.
  - Connecting the repo to RStudio.
  - Latex and Markdown in RStudio.
  - Connecting to Zotero files.

Week 03, 01/13 - 01/17: Creating composite indexes

- Factor analysis.
- Cluster analysis.
- Using plain arithmetic.

Week 03, 01/20 - 01/24: Tabular Data: univariate categorical case

- Visualizing Nominal data.
- Visualizing Ordinal data.

Week 04, 01/27 - 01/31: Tabular Data: univariate numerical case

- Visualizing Numerical data.

Week 05, 02/03 - 02/07: Tabular Data: bivariate case

- Plotting categorical-categorical associations.
- Plotting categorical-numerical relationships.
- Plotting numerical-numerical relationships.

Week 06, 02/10 - 02/14: Spatial Data

- Introducing shapefiles: point, line and polygon. Projection and coordinate system.
- Plotting multiple layers of spatial data.

Week 07, 02/17 - 02/21: Introduction to Python

- Data frames in Pandas.
- Getting text data from Twitter.

Week 08, 02/24 - 02/28: Network Data

- Plotting network elements.
- Network layouts.

Week 09, 03/02 - 03/06: Visualizing Textual Data in Python

- Visualizing text and word clouds.
- Visualizing sentiment.
- Visualizing topics.

Week 10, 03/09 - 03/13: Group ORAL Presentation.

Week 11, 03/16 - 03/20: Final week: submitting final report.

About religious accommodations

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at Religious Accommodations Policy[1]. Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form[2].

[1] https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/
[2] https://registrar.washington.edu/students/religious-accommodations-request/