



## STAT3622 Data Visualization (2019-20 Semester 2)

### Course Outline

<b>Instructor:</b>	Dr. Aijun Zhang
<b>Email:</b>	ajzhang@hku.hk
<b>Office:</b>	RR224
<b>Telephone:</b>	3917 1984
<b>Lecture Hours:</b>	Tuesday 1:30pm – 4:20pm (T5)
<b>Tutor:</b>	Yifeng Guo (gyf9712@hku.hk; RR114)
<b>Tutorial Hours:</b>	TBD

#### **Course Objectives:**

This course will focus on exploratory data analysis with statistical graphics and interactive data visualization. Students will learn how to display, communicate and analyze data, using a set of packages based on R and Python programming.

#### **Prerequisites:**

STAT2602 (Probability and Statistics II) or STAT3902 (Statistical Models).

#### **Contents and Topics:**

Data science, data manipulation, exploratory data analysis, statistical graphics, interactive data visualization, shiny applications, selected case studies.

#### **Intended Learning Outcomes:**

- Choose the best chart that fits the data
- Communicate effectively using statistical graphics
- Create compelling visualization via programming tools

#### **Assessment Method:**

Continuous:	Participation, in-class quizzes and homeworks	40%
Final Project:	DataViz app, oral presentation and written report	60%

**Course Website:** <http://stat3622.saas.hku.hk/> & <http://moodle.hku.hk/>

**Programming:** R, Python, D3.js

## References and Online Materials:

1. *R for Data Science* (2017 O'Reilly) by Golemund and Wickham. <http://r4ds.had.co.nz/>
2. *ggplot2: Elegant Graphics for Data Analysis* (2nd Edition, 2016 Springer) by Hadley Wickham. <http://had.co.nz/ggplot2/>
3. *IPython Interactive Computing and Visualization Cookbook* (2018 Packt) by Cyrille Rossant. <https://ipython-books.github.io/>
4. *D3.js in Action* (2nd Edition, 2017 Manning) by Elijah Meeks. <https://www.manning.com/books/d3js-in-action-second-edition>
5. Yau, N. (2011). *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Wiley. <http://book.flowingdata.com/>
6. RStudio Cheat Cheats. <https://www.rstudio.com/resources/cheatsheets/>

This cheat sheet, titled "Data Visualization with ggplot2 Cheat Sheet", provides a comprehensive overview of ggplot2. It is organized into several sections: Basics, Geoms (Graphical Primitives, One Variable, Two Variables, Three Variables), and Continuous Functions. Each section includes code snippets for creating various plots and visualizations, such as scatter plots, histograms, and faceted plots. The cheat sheet also covers styling options like color, size, and shape, and provides tips for handling data with missing values and faceting.

This cheat sheet, titled "Interactive Web Apps with shiny Cheat Sheet", provides a comprehensive overview of Shiny. It is organized into several sections: Basics, Building an App, and Outputs. The "Building an App" section is particularly detailed, showing how to create a Shiny app by adding arguments to the `fluidPage()` and `server` functions. It includes code snippets for creating various UI elements like text inputs, sliders, and plots, and for handling data in the server function. The "Outputs" section lists various output functions like `renderText()`, `renderPlot()`, and `renderTable()`, along with their respective UI elements.