PhD Applied Data Analysis Requirement

Marina Meilä

Approved by Statistics Department May 2019

The graduate program of the Department of Statistics at the University of Washington is based on a View of Statistics which includes the following statement: “Development of useful statistical methodology cannot take place in a vacuum. To be scientifically relevant this development should be problem-driven, motivated and guided by applications of scientific importance. Identifying and understanding important applications requires interaction with other disciplines that acquire and analyze data. Collaborative research is therefore essential to the viability and growth of Statistics.” (https://www.stat.washington.edu/academics/graduate/programs/phd). With this view in mind, all Ph.D. students must satisfy an Applied Statistics Requirement consisting of two components:

1. Satisfactory performance in Stat 599 (joint with Biostat 590)
2. A collaborative Applied Data Analysis project as delineated here.

A Consulting Committee of two Statistics faculty will oversee the fulfillment of Applied Data Analysis requirement.

(a) Prerequisite to approval to carry out an Applied Data Analysis project is completion of Stat 599.

(b) The student is responsible for finding a client with a real research problem involving a substantial data analysis component. The client must work in a field outside of Statistics or Biostatistics. The aim of the project should be a substantive scientific question in that field. A black box predictive approach (for instance training a neural network for prediction) would not be adequate because the black box algorithms don’t link to the science. Similarly, a project which is predominantly computational and not inferential would not be acceptable.

(c) The project must be of larger scope and time commitment than can be addressed by students in the regular consulting class, Stat 599; i.e. equivalent to 3 credits of independent study.

(d) Students will notify the Consulting Faculty Committee when they are interested in taking on an Applied Data Analysis Project.

(e) The client cannot be a professional statistician, although a professional consulting statistician associated with the research project may be involved in a supervisory role in the same way that any other faculty member in the Statistics or Biostatistics may supervise as specified in part 2g below.

(f) The student must submit to the Consulting Committee a written project proposal detailing the scientific aims of the analysis, the source of data, and the types of analyses and statistical challenges anticipated. This proposal will typically be short (1-2 pages) and must be approved (by signature or email) by the client.

(g) The project proposal must be approved by the Consulting Committee. The student should register for 3 credits of Stat 600 with one of the committee members.
The project will require a written report on the results of the analysis. Length of the report: at most 10 pages, with an additional 3 pages allowed to accommodate figures only; the references pages are not counted towards this page limit. The report must be written to address the scientific aims and needs of the client, and in a language accessible to a non-specialist. It should generally follow the guidelines for written reports for the statistical consulting class (Stat 599), unless a different format is requested by the client.

The report should be submitted to the Consulting Committee within one quarter of the date of approval of the project. If the report is not submitted by the end of the quarter in which the student registers for STAT 600, the student will receive an Incomplete grade, which will be converted to a Pass when the project is approved.

The completed project must be approved by the Consulting Committee. Input from the client may be submitted, but the client will not have a formal vote on approval of Analysis Project.

- **Finding clients and research problems:** It is the student’s responsibility to find a suitable client and research project. A substantial number of suitable projects are likely to arise from clients of the regular consulting class (Stat 599/Biostat 590) who have data analysis projects that are too large and require too much ongoing collaboration to be handled in the consulting class. A few such projects seem to arise each quarter. However, students may find projects on their own and they are encouraged to consult with Stat, Biostat or other faculty for possible projects.

- **Paid compensation for work carried out for the Applied Data Analysis Requirement:** Students are permitted (but not required) to be paid for the work they do on this data analysis project. Paid work may be carried out either under an RA, usually in the client’s department, or for hourly wages paid through the Department of Statistics Center for Statistical Consulting. International students should consult with the University of Washington International Student Services office (iss.washington.edu) regarding opportunities and constraints on payment for work satisfying this consulting requirement.

- **Petitioning for satisfying this requirement based on other work experience:** Like any other requirement, students can petition for a waiver of this requirement based on previous or other proposed consulting work experience. If the student plans to require a waiver based on work experience or an internship initiated during their enrollment in Statistics at the U.W., the student must obtain prior approval of a submitted proposal as 2f above.

In more detail

- The project must have a substantial data analysis component.
- The proposal is accompanied by a letter/email from the client, confirming that the data analysis work (but not the data itself) can be presented in a report in order to satisfy the requirement. The client must also confirm the dates of the internship.
- The client/internship supervisor must work in a field outside Statistics OR, in the case the team comprises other statisticians, the supervisor must explain in a letter or email the role of the student in the data analysis. A project proposal will not be approved if the student played a role deemed too subordinate to other statistician.
- The proposal must be submitted for approval at least 30 days before the end of the internship.