Introduction:

* It is challenging for many new PharmD graduates is gain entry level employment because availability within community pharmacies, clinical residencies, and industry fellowships remains limited. This is a tragedy as pharmacy students are educated at an advanced level in many areas of healthcare science, including literature review, biochemistry, and pathophysiology. Knowledge in these areas is useful in clinical informatics research, although, skills pertaining specifically to data science, database management and statistical analysis are required for students to transition into job outside of traditional career paths.

Course Overview:

* Coursera offers a course centered in familiarizing students from diverse back grounds with the fundamental of data science. The course divides 300 hours of coursework into 11 weeks of material. The student must complete 5 quizzes per week before advancing to the next week and each quiz must be passed with a minimum of 80%. The final assessment of this course is a capstone project which requires the students to apply the skills they learn in a project on public health. I propose that these courses be considered for an elective APPE rotation by condensing the 11-week course into 6 weeks. This requires the student to complete between 40-50 hours per week of course material. Cost for the course would be expected of the student at $50/4 weeks or $100 for 6 weeks.

|  |  |
| --- | --- |
| Week 1 | Coursework **(43 hours total)**   * The Data Scientist’s Toolbox **(18 hours)**   + Learning Objectives     - Introduction to R and RStudio     - Introduction to GitHub and other tools     - Understanding data problems     - Create GitHub repository   + Assessments **(must pass with 80% or higher)**     - 5 quizzes * R Programming **(25 hours)**   + Learning objectives     - Understanding critical programing language concepts     - Configure statistical programming software     - Make use of R loop functions and debugging tools     - Collect detailed information using R profiler   + Assessments **(must pass with 80% or higher)**     - 5 quizzes |
| Week 2 | Coursework **(50 hours total)**   * Getting and Cleaning Data **(30 hours)**   + Learning Objectives     - Understand common data storage systems     - Use R for text and date manipulation     - Apply data cleaning basics to make data “tidy”     - Obtain usable data from the web. APIs, and databases   + Assessments **(must pass with 80% or higher)**     - 5 Quizzes * Exploring Data Analysis **(20 hours)**   + Learning Objectives     - Understand analytic graphics and the base plotting system in R     - Use advanced graphing systems such as the Lattice system     - Make graphical displays of very high dimensional data     - Apply cluster analysis techniques to locate patterns in data   + Assessments **(must pass with 80% or higher)**     - 5 Quizzes |
| Week 3 | Coursework **(50 hours total)**   * Reproducible Research **(25 hours)**   + Learning Objectives     - Organize data analysis to help make it more reproducible     - Write up a reproducible data analysis using knitr     - Determine the reproducibility of analysis project     - Publish reproducible web documents using Markdown   + Assessments **(must pass with 80% or higher)**     - 5 quizzes * Statistical Inference **(25 hours)**   + Learning Objectives     - Understand the process of drawing conclusions about populations or scientific truths from data     - Use p-values, confidence intervals, and permutation tests     - Describe variability distributions, limits, and confidence intervals     - Make informed data analysis decision   + Assessments **(must pass with 80% or higher)**     - * 5 quizzes |
| Week 4 | Coursework **(50 hours total)**   * Regression Models **(50 hours)**   + Learning Objectives     - Use regression analysis, least squares and     - Investigate analysis of residuals and variability     - Understand ANOVA and ANCOVA models cases     - Describe novel models such as scatterplots and smoothing   + Assessment **(must pass with 80% or higher)**     - 5 quizzes |
| Week 5 | Coursework **(40 hours total)**   * Practical Machine learning **(10 hours)**   + Learning Objectives     - Use the basic components of building and applying prediction functions     - Describe machine learning methods such as regression or classification tress     - Understand concepts such as training and tests sets, overfitting, and error rates     - Explain the complete process of building prediction functions   + Assessments **(must pass with 80% or higher)**     - 5 Quizzes * Developing Data Products **(30 hours)**   + Learning Objectives     - Develop basic applications and interactive graphics using GoogleVis     - Build an R Markdown presentation that includes a data visualization     - Use Leaflet to create interactive annotated maps     - Create a data product that tells a story to a mass audience   + Assessments **(must pass with 80% or higher)**     - 5 quizzes |
| Week 6 | Coursework **(40 hours)**   * Data Science Capstone **(40 hours)**   + Learning Objectives     - Create a useful data product for the public     - Build an efficient and accurate prediction model     - Apply your exploratory data analysis skills     - Produce a presentation deck to showcase your findings   + Assessments     - **Capstone Objectives** |