Week 1. Central tendency measures, q-q plots, and kernel density estimation. The Central Limit Theorem of the log scale, the lognormal distribution for real data.

Data and examples: 234,986 individual hourly wages of US workers. Is America rich? It depends on which tendency measure you use: mean, median, or mode.

Week 2. Bivariate normal density, regression as conditional mean, principal excess, 2D kernel density estimation, and nonparametric regression. Disease mapping and its application to the hotspot localization.

Data and examples: kernel density estimation for profit and sales among the 2000 Forbes' biggest companies: which company to invest in? Distribution of the GPS locations of 10,439 lung cancers in 259 towns of New Hampshire.

Week 3. Multidimensional random vectors, partial correlation, heatmap correlation graphics animation.

Data and examples: Ordinary and partial correlation heatmap analysis of returns for the 17 largest companies, and paleontology human skeleton reconstruction using Goodman osteometric data from 1,500 bone remains worldwide.

Week 4. Multiple linear regression in matrix form, Gauss-Markov theorem, vector-space geometry of the least squares, and the coefficient of determination, interpretation of regression coefficients on the log scale.

Data and examples: Sir Francis Galton's regression and human evolution.

Week 5. The Big Theorem of regression analysis, prediction by regression, autoregression, and pitfalls, the mean and individual prediction by regression, general linear hypothesis, and the F-test.

Data and examples: Gender difference in salary, prediction of AMAZON stock prices using autoregression.

Week 6. Stochastic inequality of distributions via cdfs. Receiver Operator Characteristic (ROC) curve and AUC, sensitivity, false positive, total classification error, and optimal threshold/cutoff point, binormal ROC curve.

Data and examples: comparison of salaries between Connecticut and Vermont residents, identification of heart attack patients using blood pressure measurements, and family income as a predictor of mortgage failure.

Week 7. Logistic and Poisson regressions for classification, multivariate maximum likelihood parameter estimation, Fisher information matrix, confidence band, Wald

and likelihood testing for hypothesis testing, regression confidence band, logistic regression for binary classification.

Data and examples: classification of Democrats and Republicans based on their height, identification of mortgage non-defaulters using a set of predictors, and prediction of the number of traffic tickets using a driver's information. Optimal cutoff for identifying active Amazon shoppers.

Week 8. Principal component analysis (PCA) for data dimension reduction, linear projection onto a line, and a plane. PCA for binary classification. Using PCA for object/subject ranking, projection quality, and PCA pitfalls.

Data and examples: 5,000 applicant data for admission of 1,000 college students, identification of fraudulent banknotes based on six bill measurements.

Week 9. Hierarchical cluster analysis and K-means algorithm for unsupervised learning, dendrogram, statistical model for the K-means algorithm, broken-line algorithm for the number of clusters, variance decomposition, and the coefficient of determination, testing the no-clusters fundamental hypothesis.

Data and examples: city crime classification and identification of the number of clusters in the Goodman anthropology dataset.

Week 10. Nonlinear regression, nonlinear least squares (NLS) and the geometry, Gauss-Newton algorithm, statistical properties of the NLS estimators.

Data and examples: fitting marathon data for estimation of the world's best absolute time, the free-fall equation from Aristotle through Galileo/Newton to the hammer and feather experiment on the Moon.