Data Science in Bioinformatics – 2017

Lectures: Analysis of Biological Data
Exercises: online R for Data Science + own datasets and questions

Week	Week	Lectures	Person	Exercises
PART 1. INTRODUCTION TO STATISTICS				
1	35	1. Statistics and samples 2. Displaying data	PV	R intro RfDS 1,2,3 Visualization Data: genomes, gc and gene count
2	36	3. Describing data	PV	RfDS 4,5 Data: 1000genomes, allele frequencies
3	37	4. Estimating with uncertainty	TB	RfDS 4,5 continued
4	38	5. Probability	TB	
5	39	6. Hypothesis testing 19.1 Computer-intensive methods	PV	
PART 2. PROPORTIONS AND FREQUENCIES				
6	40	7. Analyzing proportions	PV	
7	41	8. Fitting probability models to frequency data 9. Contingency analysis: associations between categorical variables	TB	
FALL BREAK – Week 42				
8	43	10. The normal distribution 11. Inference for a normal population	TB	
9	44	12. Comparing two means 13. Handling violations of assumptions	TB	
PART 3. COMPARING NUMERICAL VALUES				
10	45	14. Designing experiments 15. Comparing means of more than two groups	TB	
PART 4. REGRESSION AND CORRELATION				
11	46	16. Correlation between numerical variables 17. Regression	PV	
PART 5. MODERN STATISTICAL METHODS				
12	47	18. Multiple explanatory variables	PV	
13	48	20. Likelihood	TB	
14	49	21. Meta-analysis: combining information from multiple studies	TB	