Noah M. Jorgenson

Bioinformatics CS6643

Lab 11

dir(pp)

['N', '\_\_builtins\_\_', '\_\_doc\_\_', '\_\_file\_\_', '\_\_name\_\_', 'allele\_count', 'allele\_freq', 'csv', 'division', 'dtype\_allele', 'itertools', 'optparse', 'parseMap', 'parsePlink', 'sys', 'valid']

dir(extra\_data)

['\_\_doc\_\_', '\_\_init\_\_', '\_\_module\_\_', 'ct\_allelic', 'data', 'get\_genotype', 'maf', 'mapArray', 'pedArray', 'snp\_list', 'split', 'split\_alleles', 'total\_female', 'total\_male', 'total\_unknown\_sex']

print extra\_data.pedArray.dtype

[('Family\_ID', '|S10'), ('Individual\_ID', '|S10'), ('Paternal\_ID', '|S10'), ('Maternal\_ID', '|S10'), ('Sex', '<i4'), ('Phenotype', '<i4'), ('rs17121574', '|S2'), ('rs754238', '|S2'), ('rs11203962', '|S2'), ('rs6999231', '|S2'), ('rs17178729', '|S2'), ('rs10105623', '|S2'), ('rs2460915', '|S2'), ('rs7835221', '|S2'), ('rs2460911', '|S2'), ('rs12156420', '|S2'), ('rs17786052', '|S2'), ('rs529983', '|S2'), ('rs630969', '|S2'), ('rs2460914', '|S2'), ('rs607499', '|S2'), ('rs634228', '|S2'), ('rs556531', '|S2')]

Cool lab, this definitely helps me understand a little better what’s inside of this .ped file. Hopefully I can apply this to my project and better understand the map data. I think it’ll be helpful now that I’ve manually worked with the data to use Plink and see what it presents now that I better understand what Plink is doing behind the scenes, so to speak.