BIOL 0495

1. **Datasets (In-Class)**
    a. **Lecture 2**
    b. **Lecture 3**
    c. **Lecture 4**
    d. **Lecture 5**
    e. **Lecture 6**
    f. **Lecture 7**
    g. **Lecture 8**
    h. **Lecture 9**
    i. **Lecture 10**
2. **Datasets (Labs)**
    a. **Lab 0**
    b. **Lab 1**
3. **Interactive Jupyter Notebooks**

TEMPLATE

_____

Lecture #

Topic/Learning Outcomes:

Datasets:

Source/Reference:

Datasets (In-Class)

Lecture 2

Topic/Learning Outcomes:

At this stage of the course, topics are typically kept broad to appeal to a wider range of students and to minimize the need for domain-specific knowledge. I tried to find a dataset that focuses on population similar to “gender and faculty divisions at Brown” prior, but I could not find one because I have to request for data and it is not in the Common Datasets Brown OIR.
The goal of this dataset should teach addition rule, independence, conditional probability

Datasets:

Columns: **CountyID, State, County, Total Population, Average Income, and Commute Time**

CountyId	State	County	TotalPop	Men	Women	Hispanic	White	Black	Native	Asian	Pacific	VotingAge	Income	IncomeErr	IncomePer	In
1001	Alabama	Autauga C	55036	26899	28137	2.7	75.4	18.9	0.3	0.9	0	41016	55317	2838	27824	
1003	Alabama	Baldwin C	203360	99527	103833	4.4	83.1	9.5	0.8	0.7	0	155376	52562	1348	29364	
1005	Alabama	Barbour C	26201	13976	12225	4.2	45.7	47.8	0.2	0.6	0	20269	33368	2551	17561	
1007	Alabama	Bibb Coun	22580	12251	10329	2.4	74.6	22	0.4	0	0	17662	43404	3431	20911	
1009	Alabama	Blount Cou	57667	28490	29177	9	87.4	1.5	0.3	0.1	0	42513	47412	2630	22021	
1011	Alabama	Bullock Co	10478	5616	4862	0.3	21.6	75.6	1	0.7	0	8212	29655	5376	20856	
1013	Alabama	Butler Cou	20126	9416	10710	0.3	52.2	44.7	0.1	1.1	0	15459	36326	2701	19004	
1015	Alabama	Calhoun C	115527	55593	59934	3.6	72.7	20.4	0.2	1	0	88383	43686	1491	23638	
1017	Alabama	Chambers	33895	16320	17575	2.2	56.2	39.3	0.3	1	0	26259	37342	2011	22002	
1019	Alabama	Cherokee	25855	12862	12993	1.6	91.8	5	0.5	0.1	0	20620	40041	2316	23010	
1021	Alabama	Chilton Co	43805	21554	22251	7.7	80.4	9.5	0.4	0.4	0	31776	43501	2877	23368	
1023	Alabama	Choctaw C	13188	6277	6911	0.5	56.3	42.1	0	0.1	0	10454	32122	2797	20994	

Addition Rule: We want to find the probability of selecting a person who is either from a high-income county OR has a long commute
Independence: We want to determine if gender and commute time are independent variables across counties
Conditional Probability: We want to find the probability of a person being from a high-income county, given that they're from a county with a long average commute.

Source/Reference

ACS 2017

Lecture 3

Topic/Learning Outcomes:

The Pima Indians Diabetes Database is relevant for our class on discrete probability because it includes discrete values like glucose levels, blood pressure, and age, which can be analyzed to understand health outcomes. The third lecture will build on our understanding of discrete variables. Although some values are continuous, we can categorize them on a scale of 1 to 5 for both glucose and blood pressure levels. The dataset's outcome variable indicates whether a person has diabetes (1) or does not (0), which follows a binomial distribution. This binary classification enables us to investigate how discrete predictors affect the likelihood of developing diabetes, serving as a strong motivation to explore the role of predictors in health outcomes

Datasets:

Columns:

	A	B	C	D	E	F	G	H	I
1	Pregnancie	Glucose	BloodPres	SkinThickn	Insulin	BMI	DiabetesP	Age	Outcome
2	6	148	72	35	0	33.6	0.627	50	1
3	1	85	66	29	0	26.6	0.351	31	0
4	8	183	64	0	0	23.3	0.672	32	1
5	1	89	66	23	94	28.1	0.167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	0.201	30	0
8	3	78	50	32	88	31	0.248	26	1
9	10	115	0	0	0	35.3	0.134	29	0
10	2	197	70	45	543	30.5	0.158	53	1
11	8	125	96	0	0	0	0.232	54	1
12	4	110	92	0	0	37.6	0.191	30	0
13	10	168	74	0	0	38	0.537	34	1

Source/Reference:

[Pima Indian Datasets]

Lecture 4

Topic/Learning Outcomes:

Datasets:

Source/Reference:

Lecture 5

Topic/Learning Outcomes:

Datasets:

Source/Reference:

Lecture 6

Topic/Learning Outcomes:

Datasets:

Source/Reference:

Lecture 7

Topic/Learning Outcomes:

Datasets:

Source/Reference:

Datasets (Lab)