

3011979 Practical Python for Data Sciences and Machine Learning

L2: Python for data science and visualization

Jan 21st, 2022



Sira Sriswasdi, Ph.D.

Research Affairs, Faculty of Medicine
Chulalongkorn University

Datasets

Dataset overview

- Dialysis monitoring
- Mutations in Myelodysplastic Syndrome (MDS) patients
- Psoas muscle strength and MRI features
- Radiotherapy treatment outcome in liver cancer cases
- Absorbance spectrum of rice leaf under nutrient deprivation
- Clinical factors of refractory thyroid cancer
- Montreal Cognitive Assessment (MoCA) test scores
- Fetal doppler at various gestation ages
- Mass spectra of saliva samples from COVID-19 patient
- COVID-19 home isolation monitoring
- COVID-19 EID data

Dialysis monitoring

Patient ID	Treatment Date	Time	SBP	DBP	MAP	HR	UFR
N001	2018/09/04	2:02:00 PM	162	57	103	75	725
N001	2018/09/04	2:23:00 PM	150	58	81	73	797
N001	2018/09/04	2:56:00 PM	134	52	75	77	725
N001	2018/09/04	4:01:00 PM	135	61	90	76	725
N001	2018/09/04	5:02:00 PM	138	57	85	71	532
N001	2018/09/04	5:24:00 PM	99	82	92	65	5
N001	2018/09/04	5:25:00 PM	113	46	73	65	5
N001	2018/09/04	6:33:00 PM	155	59	79	67	5
N001	2018/09/04	6:37:00 PM	163	55	85	66	0
N001	2018/09/08	7:37:00 AM	149	70	102	71	200
N001	2018/09/08	7:43:00 AM	153	60	108	73	276
N001	2018/09/08	8:40:00 AM	140	62	97	65	251
N001	2018/09/08	9:42:00 AM	128	63	96	63	251
N001	2018/09/08	10:29:00 AM	140	62	100	62	259
N001	2018/09/08	11:25:00 AM	160	70	114	60	284
N001	2018/09/08	11:43:00 AM	176	69	124	59	259
N001	2018/09/08	11:45:00 AM	162	68	122	59	259

Mutation in MDS

Sample ID	Ensembl ID	Gene Symbol	Ref Allele	Alt Allele	Predicted Impact	Genotype	Variant Type	VAF	Population AF
T267.T267_S3	ENSG00000120952	PRAMEF2	C	T	MODERATE	hetero	missense_variant	0.5215	0.1336
T267.T267_S3	ENSG00000120952	PRAMEF2	T	C	MODERATE	hetero	missense_variant	0.5927	0.4099
T267.T267_S3	ENSG00000120952	PRAMEF2	G	A	MODERATE	hetero	missense_variant	0.3866	0.2714
T267.T267_S3	ENSG00000120952	PRAMEF2	T	G	MODERATE	hetero	missense_variant	0.5667	0.1929
T267.T267_S3	ENSG00000120952	PRAMEF2	C	T	MODERATE	hetero	missense_variant	0.1852	
T267.T267_S3	ENSG00000107890	ANKRD26	A	T	MODERATE	hetero	missense_variant	0.2255	0.3173
T267.T267_S3	ENSG00000107890	ANKRD26	C	T	MODERATE	hetero	missense_variant	0.517	0.3173
T267.T267_S3	ENSG00000107890	ANKRD26	T	C	MODERATE	hetero	missense_variant	0.5682	0.06709
T267.T267_S3	ENSG00000107890	ANKRD26	T	C	MODERATE	homo	missense_variant	0.9965	0.9447
T267.T267_S3	ENSG00000175591	P2RY2	C	T	MODERATE	homo	missense_variant	0.9967	
T267.T267_S3	ENSG00000175591	P2RY2	G	C	MODERATE	homo	missense_variant	0.9957	0.2724
T267.T267_S3	ENSG00000149308	NPAT	G	T	MODERATE	hetero	missense_variant	0.5098	0.04253
T267.T267_S3	ENSG00000149308	NPAT	C	T	MODERATE	hetero	missense_variant	0.4819	0.6156
T267.T267_S3	ENSG00000149311	ATM	A	G	MODERATE	homo	missense_variant	1	1

Psoas muscle

Patient ID	Sex	Age	Ht	Wt	Rt/Lt	Peak Force First	Peak Force Second	Duration First	Duration Second	Psoas Area Rt	Psoas Area Lt	Psoas Thickness Rt	Psoas Thickness Lt	Psoas Width Rt	Psoas Width Lt	% fat Rt	% fat Lt
N001	M	55	1.70	74	Rt	18.5	21.9	4.6	6.6	18.99	21.21	4.91	5.08	5.34	5.68	4.80	5.18
N002	F	65	1.60	72	Rt	5.5	5.0	3.8	5.2	9.58	10.21	3.44	2.85	3.80	4.96	7.68	14.80
N003	F	77	1.56	63	Rt	3.4	4.5	3.8	6.0	8.06	9.99	4.68	4.62	2.48	3.37	15.12	22.37
N004	M	57	1.60	68	Lt	7.5	14.0	5.6	4.0	18.07	17.74	4.61	5.11	5.34	4.86	6.36	4.94

Radiotherapy treatment outcome

patient ID	age	sex	liver volume	ast_b	alt_b	tb_b	hbv_b	hcv_b	cp base	ALBI grade	MLD TPS	ast_w	alt_w	tb_w	cp worst	ALBI worst grade
N001	50	0	402.3	99	18	0.7	1	0	5	2	11.046	343	21	1.34	6	2
N002	58	1	1358.4	34	321	3.92	0	0	9	2		28	120	4.08	12	3
N003	62	1	1060	28	19	0.86	0	1	5	1	6.455	30	26	0.74	5	1
N004	59	1	1613.7	87	37	1.01	0	1	6	2	17.468	242	64	0.87	7	2
N005	82	0	654.3	139	74	1.12	0	0	6	3	17.42	116	74	1.04	7	2
N006	60	1	1484.3	131	73	1.57	0	1	7	3	9.917	71	33	2.13	8	3
N007	58	1	1516	71	87	1.07	0	1	6	2		114	86	0.95	8	3
N008	85	1	1053	94	97	1.67	0	0	6	2	17.776	67	28	1.01	6	2
N009	60	1	1456.4	38	36	0.37	1	0	5	1	16.073	65	55	0.45	5	1
N010	56	1	875.8	43	32	3.02	0	1	8	2	21.184	86	51	24.76	12	3
N011	49	0	830.1	63	47	0.46	1	0	6	2	19.583	44	38	0.67	5	1

Rice leaf absorbance spectrum

Condition ID	Cultivar ID	Pi content	324 nm	334 nm	344 nm	354 nm	364 nm	374 nm	384 nm	394 nm	404 nm
1	1	86.22368	0.136531	0.077077	0.075083	0.061408	0.058935	0.054756	0.055509	0.049247	0.049859
1	1	104.7929	0.128571	0.092857	0.084086	0.05536	0.069373	0.05032	0.048282	0.049079	0.052317
1	1	167.3359	0.094961	0.128649	0.123823	0.066028	0.054652	0.049103	0.044307	0.04948	0.05518
1	1	215.9113	0.130802	0.08764	0.102281	0.085255	0.074521	0.064555	0.057626	0.057752	0.05699
1	1	223.6911	0.125862	0.089796	0.075459	0.09519	0.061393	0.070902	0.061411	0.057944	0.06273
1	1	357.4742	0.184109	0.100109	0.114789	0.054659	0.075955	0.06523	0.053606	0.059144	0.058108
1	2	220.6559	0.175105	0.129213	0.075055	0.073995	0.082015	0.069038	0.065858	0.058728	0.052119
1	2	222.6972	0.116279	0.101622	0.081101	0.061235	0.062995	0.056748	0.059075	0.056525	0.052644
1	2	230.2382	0.224806	0.173014	0.083099	0.086413	0.075548	0.073675	0.06254	0.061067	0.056556
1	2	237.8297	0.07551	0.077551	0.067408	0.072471	0.059125	0.055265	0.059878	0.052891	0.054232
1	2	305.2518	0.062069	0.058163	0.048266	0.055611	0.043684	0.042825	0.041079	0.048024	0.048737

Refractory thyroid

patient ID	age	Tsize	Tfocal	Invasion	LNmets	LNsize	dose1	tg1	antitg1	ft41	tsh1	TBS mets	Site mets	Final N
N001	55	2.7	2	0	0	0	100	4.93	368.4		100	0	0	0
N002	71			0	1	1	150	87.85	13.28	0.226	90.18	0	0	0
N003	62	3.3	1	0	0	0	100	29.8	10.2	0.4	72.8	0	0	0
N004	64	3.5	2	0	0	0	30	2.82	25.8	0.25	89	0	0	0
N005	81	1	2	1	0	0	150	51.6	10	0.4	37.72	0	0	0
N006	74	0.6	1	0	1	1.3	150	8.67	24.06	1.58	53	0	0	0
N007	76	7	1	1	0	0	150	0.04	582.9	0.22	34.7	0	0	0
N008	85	5	1	1	0	0	150	2.47	26.84	0.43	53.9	0	0	0
N009	56	2.8	1	1	0	0	100	0.5	19.32	0.4	100	0	0	0
N010	65	6.3	1	1	0	0	150	500	24.36	0.8	4.63	1	4	1
N011	59	0.5	1	1	0	0	100	0.04	43.2	0.42	41.9	0	0	0
N012	68	0.7	1	1	0	0	100	0.04	241.9	0.4	64.33	0	0	0
N013	82	3.7	2	0	0	0	100					1	2	1
N014	79	1.3	2	1	0	0	30	0.04	70.9	0.07	100	0	0	0

Montreal cognitive assessment test

ID	RID	SITEID	TRAILS	CUBE	CLOCK CON	IMMT1 W1	IMMT1 W2	DIGFO R	DIGBA CK	LETTE RS	SERIAL 1	REPEA T1	FFLUE NCY	ABSTR AN	ABSME AS
4	292	24	1	1	1	1	1	1	1	0	1	0	5	1	1
6	1326	12	0	0	1	0	0	1	1	1	1	1	7	1	0
8	376	17	1	1	1	0	1	1	1	0	1	1	5	0	0
10	1117	12	0	0	0	1	1	1	1	2	0	1	11	1	0
12	618	21	1	1	1	1	1	1	1	0	1	1	12	1	1
14	2002	48	1	1	1	1	1	1	1	1	1	1	30	1	1
16	2007	12	1	0	1	1	1	1	1	1	1	1	20	0	1
18	2010	1	1	0	1	1	1	1	1	0	1	1	18	1	1
20	622	47	1	0	1	1	1	1	1	0	1	1	10	1	1
22	685	1	1	0	1	1	1	1	1	1	1	1	11	1	1
24	2011	48	1	1	1	1	1	1	1	0	1	1	22	1	1
26	729	1	1	0	1	1	1	1	1	0	1	1	9	1	1
28	548	11	1	1	1	1	1	1	0	0	1	1	13	0	0

Fetal doppler

GA	Age	Weight	Height	Gravid	Para	FHR	AT R	AT L	ET R	ET L	ET2 R	ET2 L	ET3 R	ET3 L
20	42	80	152	2	1	156	22	20	160.2	168.8	190.2	195.6	200.4	217.8
20	15	48.1	156	1	0	145	21	26.5	166.7	167.25	175.6	180.55	193.3	200
20	38	47.3	143	1	0	140		22		147.25		161.7		177.25
20	28	51.2	155	1	0	157		27		160		172		187
20	24	71.9	156	2	1	156	21		156		170		186	
20	32	64.3	156	2	1	160	22		171		180		191	
20	26	55	152	3	2	138	24.5		185.55		200		225.55	
20	19	46.4	153	1	0	135	20		163.33		177.78		204.44	
20	30	61	149	2	1	147	28		188.89		198.89		212.22	
20	34	68.1	150	1	0	142		22		174.44		191.11		208.89
20	18	50	155	1	0	145		21		158.89		172.22		184.44
20	24	49.8	156	1	0	148	27		155.56		168.89		183.33	
20	28	43.7	150	1	0	149		28		156.67		171.11		182.22
20	35	52.5	160	3	1	159	28		156.67		167.78		176.67	
20	27	50.8	150	2	1	150	24		175.56		188.89		213.33	

Mass spectra of saliva

Patient ID	PCR result	Manual coloring	306	313	331	341	351	358	365	375	395	411	436
2650	pos	red		2.578234		1.955865				4.905971	2.790664	2.784736	2.713982
2650	pos	yellow	0.318931	2.444237	3.968141			3.100324		4.804096	1.892374	2.05817	1.964251
2650	pos	orange			3.958593	1.500287		3.004807		4.639372	2.087643	2.134218	
2650	pos	yellow	1.034389	2.457055	3.90205	1.382046		3.045046	1.261373	4.702395	2.474538	2.343331	
2650	pos	orange		2.455102	3.759621	1.36772		2.56416		4.401254	2.250659	2.120021	
2650	pos	yellow		2.675821	4.061289	1.667493		3.082155		4.650782	2.180814	2.08955	
2650	pos	orange		2.15872		1.509767		3.145033		4.438086	2.408704	2.261207	2.925485
2651	pos	green	1.731657	3.339431	4.908488	2.301817	1.595719	4.366962		5.574608	4.074637	4.185921	3.97063
2651	pos	green	1.661588	3.271239	4.817928	2.155788		4.360285		5.358648	4.233133	4.338697	3.806994
2651	pos	green	1.879286	3.486402		2.297723		4.523086		5.546462	4.579405	4.472148	4.023882
2651	pos	green	2.048947	3.463339	4.892583	2.234982		4.428339		5.505815	4.554893	4.576874	3.923881

Home isolation

patient ID	date	time	gender	wt	ht	body temp	pre SpO2	post spO2	pre pulse	post pulse	pre dyspnea	post dyspnea
11937884	7/9/21	9:58:00	female	58	157	35.7	98	98	75	88	1	2
11937884	7/10/21	1:16:00	female	58	157	36	97	96	78	84	1	2
11937884	7/10/21	10:10:00	female	58	157	36.1	98	98	78	82	1	1
11937884	7/11/21	1:16:00	female	58	157	35.6	95	97	76	78	0	1
11937884	7/11/21	10:14:00	female	58	157	36.1	98	99	84	102	1	2
11937884	7/12/21	1:08:00	female	58	157	35.9	99	99	78	104	0	2
11937884	7/12/21	10:04:00	female	58	157	36.7	99	99	94	104	1	2
11937884	7/13/21	1:03:00	female	58	157	35.9	97	99	97	101	1	2
11937884	7/13/21	6:23:00	female	58	157	35	98	None	88	None	0	None
11937884	7/13/21	10:24:22	female	58	157	36.8	99	99	87	107	1	2
11937884	7/14/21	1:09:41	female	58	157	35.6	97	99	70	106	1	2
11937884	7/14/21	1:17:22	female	58	157	35.6	97	99	70	106	1	3
11937884	7/14/21	9:37:33	female	58	157	36.2	97	99	77	101	1	2
11937884	7/15/21	1:13:28	female	58	157	35.7	99	99	82	101	1	3
11937884	7/15/21	10:06:16	female	58	157	36	99	99	85	107	1	3
11937884	7/16/21	0:59:55	female	58	157	35.7	99	99	87	108	1	3
11937884	7/16/21	10:05:54	female	58	157	35.4	99	99	98	110	1	3
11937884	7/17/21	1:04:21	female	58	157	36.1	99	99	98	109	1	3
11937884	7/17/21	10:05:15	female	58	157	35.9	99	99	98	115	1	3
11937884	7/18/21	1:04:23	female	58	157	36.1	97	98	100	112	1	3
11937884	7/18/21	10:07:18	female	58	157	36.6	99	99	99	115	1	3
11937884	7/19/21	1:04:46	female	58	157	36.2	99	99	96	110	1	3
11937884	7/19/21	10:04:38	female	58	157	35.6	97	99	100	115	2	3
11937884	7/20/21	1:02:28	female	58	157	36.2	96	99	97	116	1	3
11937884	7/20/21	10:27:54	female	58	157	36.2	98	99	92	115	1	3

EID COVID-19

patient id	sex	age	bmi	hct	platelet	wbc	lymphocyte	Cr	tot bili	AST	ALT	ORF-1ab CT	supplement O2	FAVI	remdes	LOS
N001	F	32	19.56	36	257000	8460	4187	0.52	0.49	16	23	18.86	0	1	0	7
N002	M	55	22.58	38.2	128000	4780	960	1.27	0.69	30	42	21.2	0	1	0	14
N003	F	34	19.02	36.7	397000	4720	1520	0.66	0.47	73	38	23.25	0	1	0	3
N004	F	74	22.35	36.8	208000	4130	1320	0.68	0.27	19	34	15.02	0	1	0	13
N005	M	49	23.8	39.3	278000	3720	1380	0.65	0.62	28	58	14.81	0	0	0	6
N006	F	56		38.3	180000	3710	890	0.8	0.49	15	23	25.05	0	1	0	13
N008	F	29	23.38	41.1	286000	3920	1020	0.66	0.37	23	21	31.31	0	0	0	5
N009	M	37	25	46.1	154000	4710	800	0.9	0.58	27	26	35	0	1	0	4
N010	F	25	18.71	32.9	179000	5940	2450	0.44	0.51	29	27	24.64	0	1	0	3
N011	F	51	25.81	34.9	287000	7060	2380	0.53	0.34	29	29	23.02	0	1	0	4
N013	F	53	20.47	31.2	292000	4360	2000	0.69	0.24	10	21	32.64	0	0	0	12

Exploratory data analysis

Some places to start

- What patterns are expected given domain knowledge?
 - Do the data agree or disagree?
- Feature analysis
 - Distribution of feature values: unimodal or multi-modal
 - Differences in feature value across classes
 - Correlation between feature and outcome
 - Correlation between features
- Multivariate analysis
 - How are samples distributed?
 - Does any feature explain the distribution?

What patterns do you see in the data?

Patient ID	Treatment Date	Time	SBP	DBP	MAP	HR	UFR
N001	2018/09/04	2:02:00 PM	162	57	103	75	725
N001	2018/09/04	2:23:00 PM	150	58	81	73	797
N001	2018/09/04	2:56:00 PM	134	52	75	77	725
N001	2018/09/04	4:01:00 PM	135	61	90	76	725
N001	2018/09/04	5:02:00 PM	138	57	85	71	532
N001	2018/09/04	5:24:00 PM	99	82	92	65	5
N001	2018/09/04	5:25:00 PM	113	46	73	65	5
N001	2018/09/04	6:33:00 PM	155	59	79	67	5
N001	2018/09/04	6:37:00 PM	163	55	85	66	0
N001	2018/09/08	7:37:00 AM	149	70	102	71	200
N001	2018/09/08	7:43:00 AM	153	60	108	73	276
N001	2018/09/08	8:40:00 AM	140	62	97	65	251
N001	2018/09/08	9:42:00 AM	128	63	96	63	251
N001	2018/09/08	10:29:00 AM	140	62	100	62	259
N001	2018/09/08	11:25:00 AM	160	70	114	60	284
N001	2018/09/08	11:43:00 AM	176	69	124	59	259
N001	2018/09/08	11:45:00 AM	162	68	122	59	259

What may explain longer length of stay?

patient id	sex	age	bmi	hct	platelet	wbc	lymphocyte	Cr	tot bili	AST	ALT	ORF-1ab CT	supplement O2	FAVI	remdes	LOS
N001	F	32	19.56	36	257000	8460	4187	0.52	0.49	16	23	18.86	0	1	0	7
N002	M	55	22.58	38.2	128000	4780	960	1.27	0.69	30	42	21.2	0	1	0	14
N003	F	34	19.02	36.7	397000	4720	1520	0.66	0.47	73	38	23.25	0	1	0	3
N004	F	74	22.35	36.8	208000	4130	1320	0.68	0.27	19	34	15.02	0	1	0	13
N005	M	49	23.8	39.3	278000	3720	1380	0.65	0.62	28	58	14.81	0	0	0	6
N006	F	56		38.3	180000	3710	890	0.8	0.49	15	23	25.05	0	1	0	13
N008	F	29	23.38	41.1	286000	3920	1020	0.66	0.37	23	21	31.31	0	0	0	5
N009	M	37	25	46.1	154000	4710	800	0.9	0.58	27	26	35	0	1	0	4
N010	F	25	18.71	32.9	179000	5940	2450	0.44	0.51	29	27	24.64	0	1	0	3
N011	F	51	25.81	34.9	287000	7060	2380	0.53	0.34	29	29	23.02	0	1	0	4
N013	F	53	20.47	31.2	292000	4360	2000	0.69	0.24	10	21	32.64	0	0	0	12

Is there any correlation between features?

Patient ID	Sex	Age	Ht	Wt	Rt/Lt	Peak Force First	Peak Force Second	Duration First	Duration Second	Psoas Area Rt	Psoas Area Lt	Psoas Thickness Rt	Psoas Thickness Lt	Psoas Width Rt	Psoas Width Lt	% fat Rt	% fat Lt
N001	M	55	1.70	74	Rt	18.5	21.9	4.6	6.6	18.99	21.21	4.91	5.08	5.34	5.68	4.80	5.18
N002	F	65	1.60	72	Rt	5.5	5.0	3.8	5.2	9.58	10.21	3.44	2.85	3.80	4.96	7.68	14.80
N003	F	77	1.56	63	Rt	3.4	4.5	3.8	6.0	8.06	9.99	4.68	4.62	2.48	3.37	15.12	22.37
N004	M	57	1.60	68	Lt	7.5	14.0	5.6	4.0	18.07	17.74	4.61	5.11	5.34	4.86	6.36	4.94

How to handle pre/post measurements?

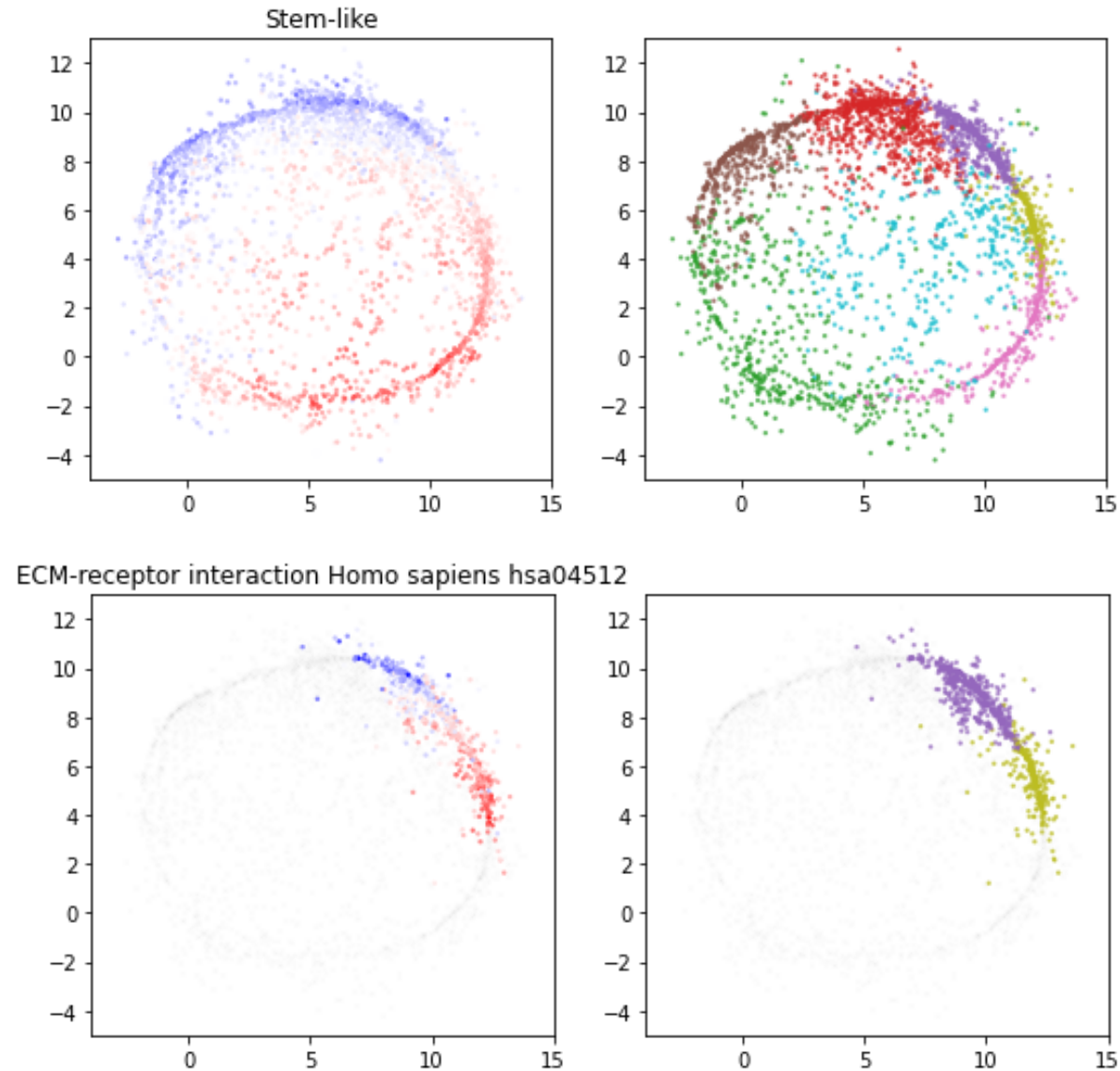
patient ID	date	time	gender	wt	ht	body temp	pre SpO2	post spO2	pre pulse	post pulse	pre dyspnea	post dyspnea
11937884	7/9/21	9:58:00	female	58	157	35.7	98	98	75	88	1	2
11937884	7/10/21	1:16:00	female	58	157	36	97	96	78	84	1	2
11937884	7/10/21	10:10:00	female	58	157	36.1	98	98	78	82	1	1
11937884	7/11/21	1:16:00	female	58	157	35.6	95	97	76	78	0	1
11937884	7/11/21	10:14:00	female	58	157	36.1	98	99	84	102	1	2
11937884	7/12/21	1:08:00	female	58	157	35.9	99	99	78	104	0	2
11937884	7/12/21	10:04:00	female	58	157	36.7	99	99	94	104	1	2
11937884	7/13/21	1:03:00	female	58	157	35.9	97	99	97	101	1	2
11937884	7/13/21	10:24:22	female	58	157	36.8	99	99	87	107	1	2
11937884	7/14/21	1:09:41	female	58	157	35.6	97	99	70	106	1	2
11937884	7/14/21	1:17:22	female	58	157	35.6	97	99	70	106	1	3
11937884	7/14/21	9:37:33	female	58	157	36.2	97	99	77	101	1	2
11937884	7/15/21	1:13:28	female	58	157	35.7	99	99	82	101	1	3
11937884	7/15/21	10:06:16	female	58	157	36	99	99	85	107	1	3
11937884	7/16/21	0:59:55	female	58	157	35.7	99	99	87	108	1	3
11937884	7/16/21	10:05:54	female	58	157	35.4	99	99	98	110	1	3
11937884	7/17/21	1:04:21	female	58	157	36.1	99	99	98	109	1	3
11937884	7/17/21	10:05:15	female	58	157	35.9	99	99	98	115	1	3
11937884	7/18/21	1:04:23	female	58	157	36.1	97	98	100	112	1	3
11937884	7/18/21	10:07:18	female	58	157	36.6	99	99	99	115	1	3
11937884	7/19/21	1:04:46	female	58	157	36.2	99	99	96	110	1	3
11937884	7/19/21	10:04:38	female	58	157	35.6	97	99	100	115	2	3
11937884	7/20/21	1:02:28	female	58	157	36.2	96	99	97	116	1	3
11937884	7/20/21	10:27:54	female	58	157	36.2	98	99	92	115	1	3

Goals of visualization

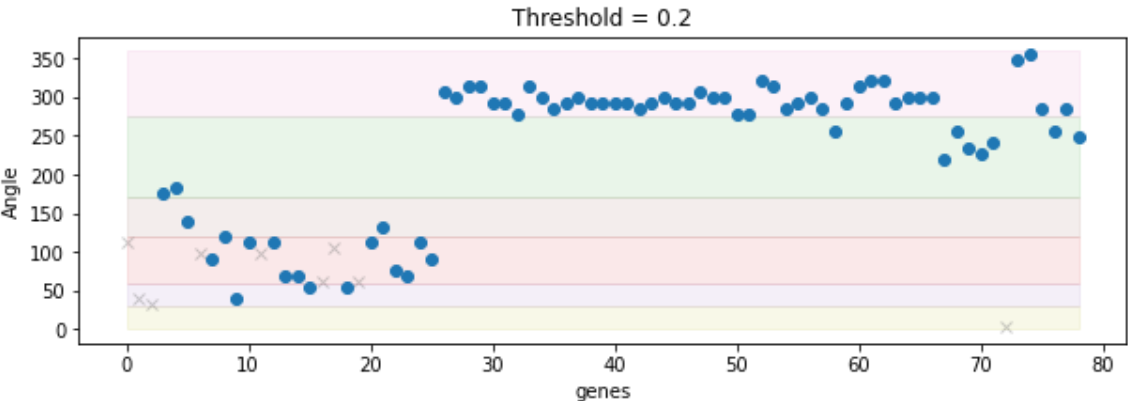
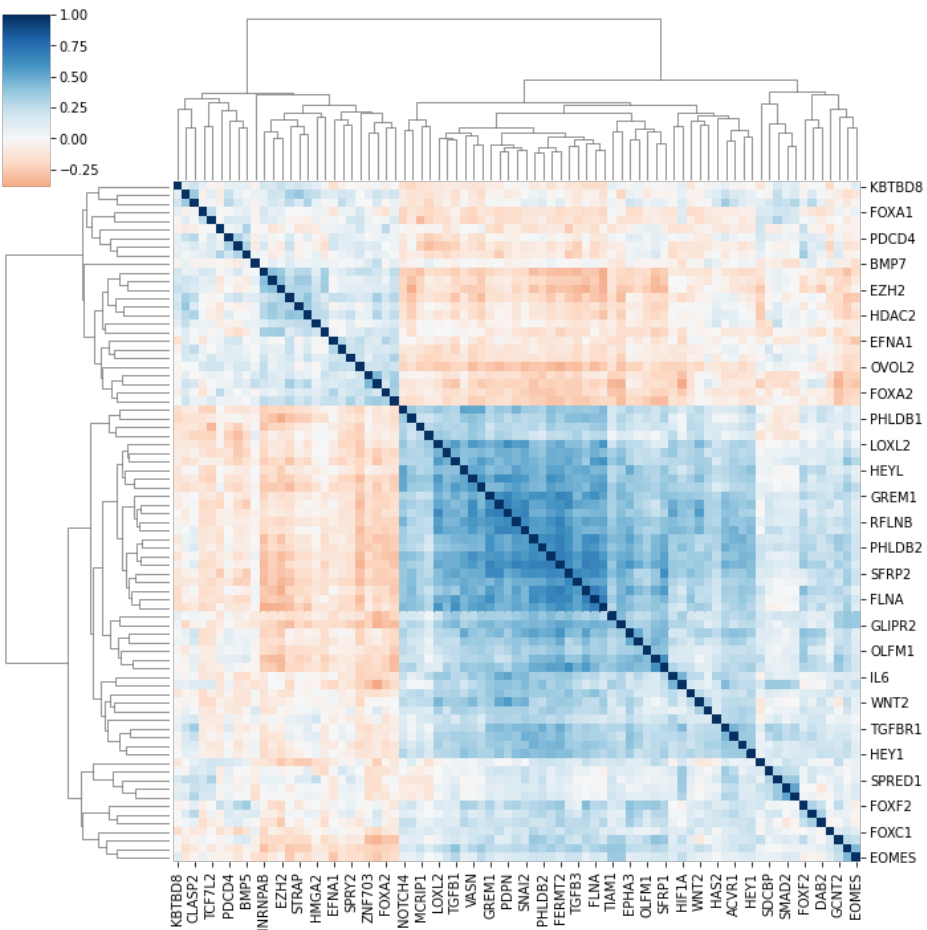
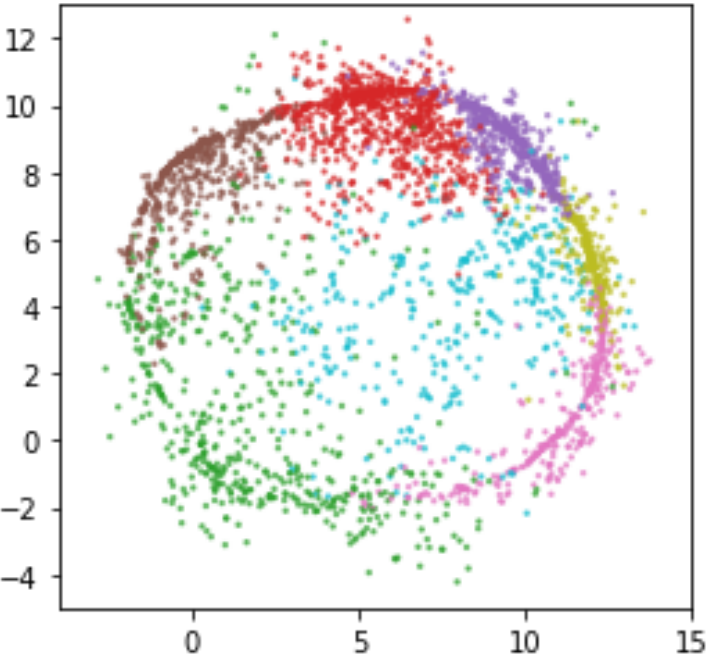
- For yourself:
 - Identify anomaly and outlier
 - Develop hypothesis
 - Understand feature
 - Error analysis

- For others:
 - Keep things reproducible
 - Always set random state!!
 - Figure should be understandable by itself + one sentence
 - Try changing plot style

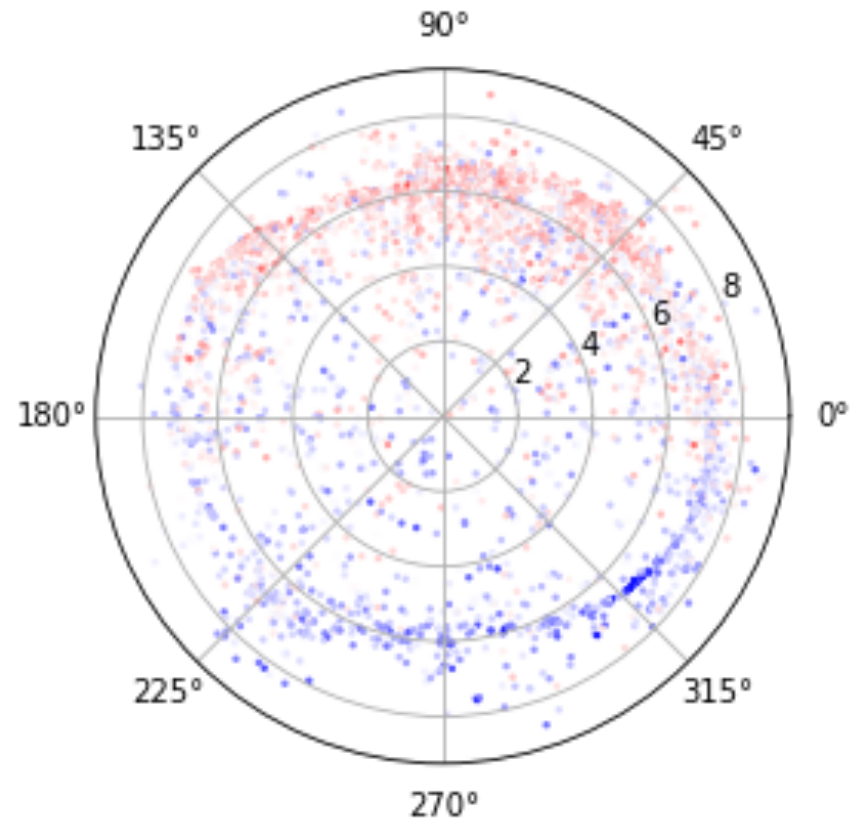
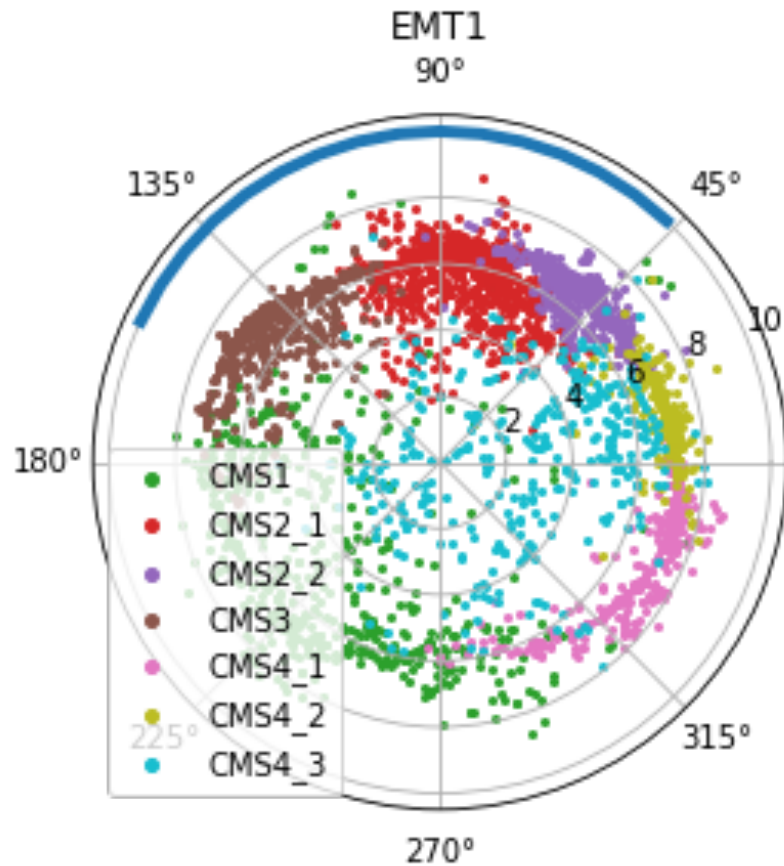
Focusing on specific comparison



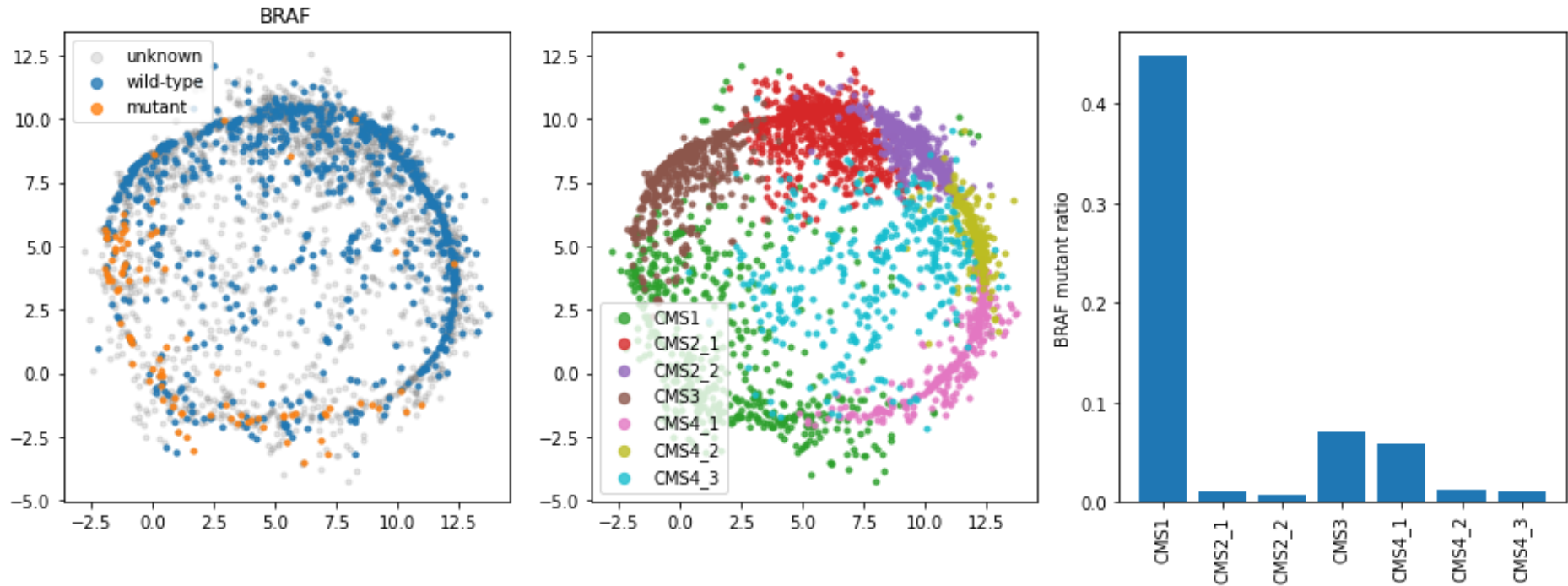
Combine plots #1



Combine plots #2

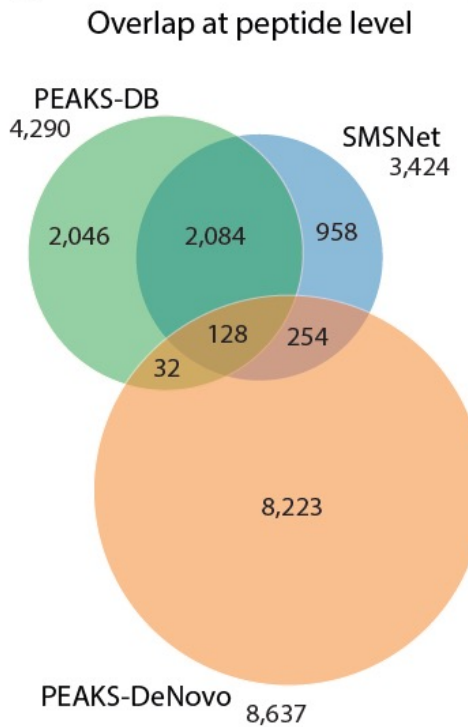


Combine plots #3

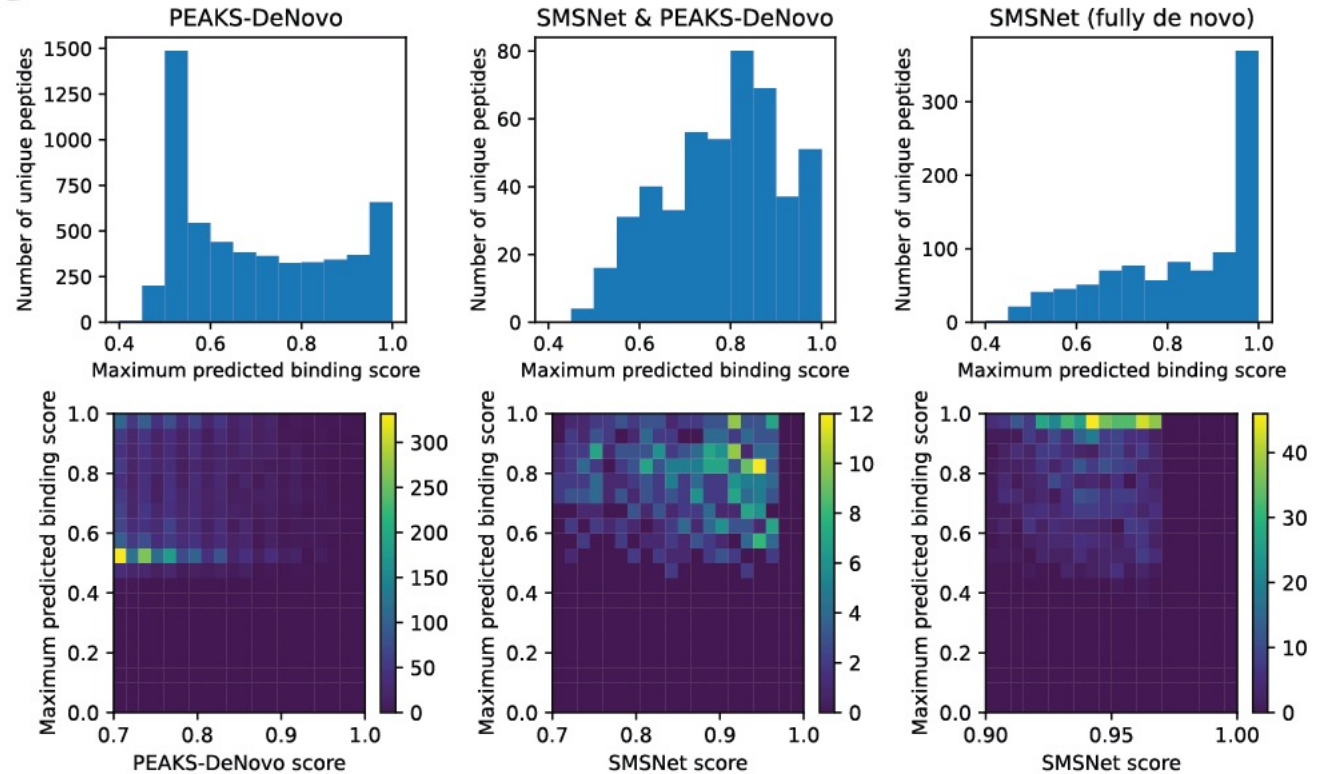


Combine plots #4

a



c



Setting up Jupyter

- Use **pip install PACKAGE** to install packages
- Running **jupyter notebook** in your Windows Command Prompt or Mac Terminal will launch the interface
- Use **jupyter notebook --notebook-dir="PATH"** to specify starting folder (where your notebook files are)
 - Replace PATH with a target directory such as:
 - C:\Courses\Spring 2022\Special Topics in Medical Physics
 - /Users/kuroneko/Documents/notebook
 - Keep the “” symbol to allow **space** in PATH name

Let's code!

- Handling data frame in Python with **pandas**
- Visualizations with **matplotlib** and **seaborn**