1. Use LASSO regressions to create the network of symptoms and COVID-19 diagnosis. Remove equations that explain less than 10% of variation in the response variables. Remove coefficients where the absolute value of coefficients is equal or less than 0.05. Remove cycles, none should exist if you always regressed response variables on independent variables that occur prior to it. From the network calculate the following

a. What is the order of occurrences of the symptoms, age, gender, and results of COVID-19 laboratory tests?

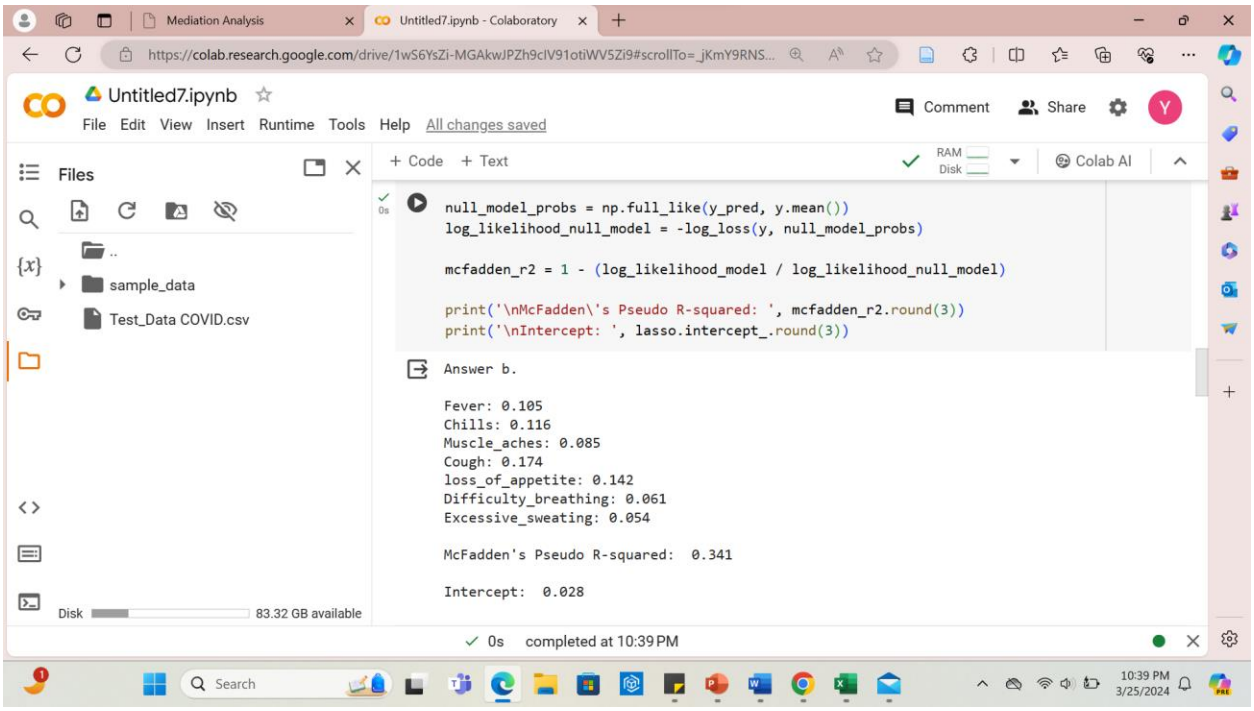
Answer:

The Order of Occurrences of the Symptoms, Age, Gender, and Results of COVID-19 Laboratory Tests is Age, Female, Shivering, Fatigue, Loss of Taste, Fever, Headaches, Loss of Smell, Chills, Muscle Aches, Diarrhea, Cough, Shortness of Breath, Runny Nose, Sore Throat, Loss of Balance, Vomiting, Joint Pain, Loss of Appetite, Wheezing, Difficulty Breathing, Excessive Sweating, and COVID-19 Test Results.

b. What are the direct predictors of COVID-19 Laboratory test results? Assume the following order for the variables: D1: Age, D2: Female, X1: Shivering, X2: Fatigue, X3: Loss of taste, X4: Fever, X5: Headaches, X6: Loss of smell, X7: Chills, X8: Muscle aches X9: Diarrhea, X10: Cough, X11: Shortness of breath, X12: Runny nose, X13: Sore throat, X14: Loss of balance, X15: Vomiting, X16: Joint pain, X17: Loss of appetite, X18: Wheezing, X19: Difficulty breathing, X20: Excessive sweating, Y: COVID-19 Test Results.

Answer:

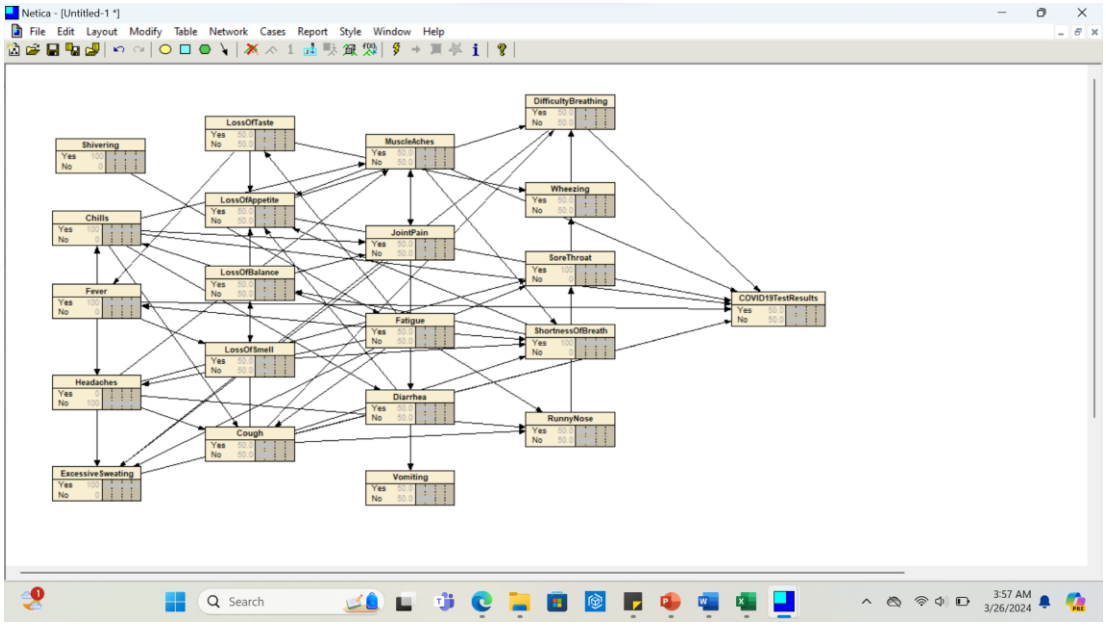
Image 1: The Direct Predictors of COVID-19 Laboratory Test Results are:



c. What is the best network that fits the data? Establish the structure of the network ignoring regressions that explain less than 10% of the variation in test results and ignoring variables where absolute value of coefficients are less than or equal to 0.05.

Answer:

Image 2: The below Netica Network was structured based on the LASSO Regression Coefficients and Intercepts:



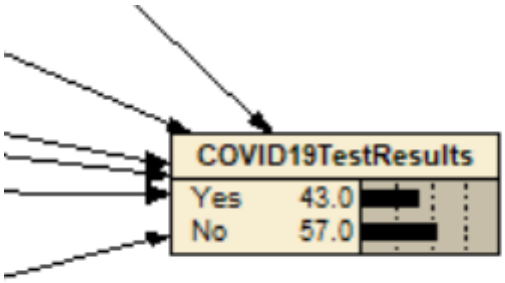
Regression Coefficients Table:

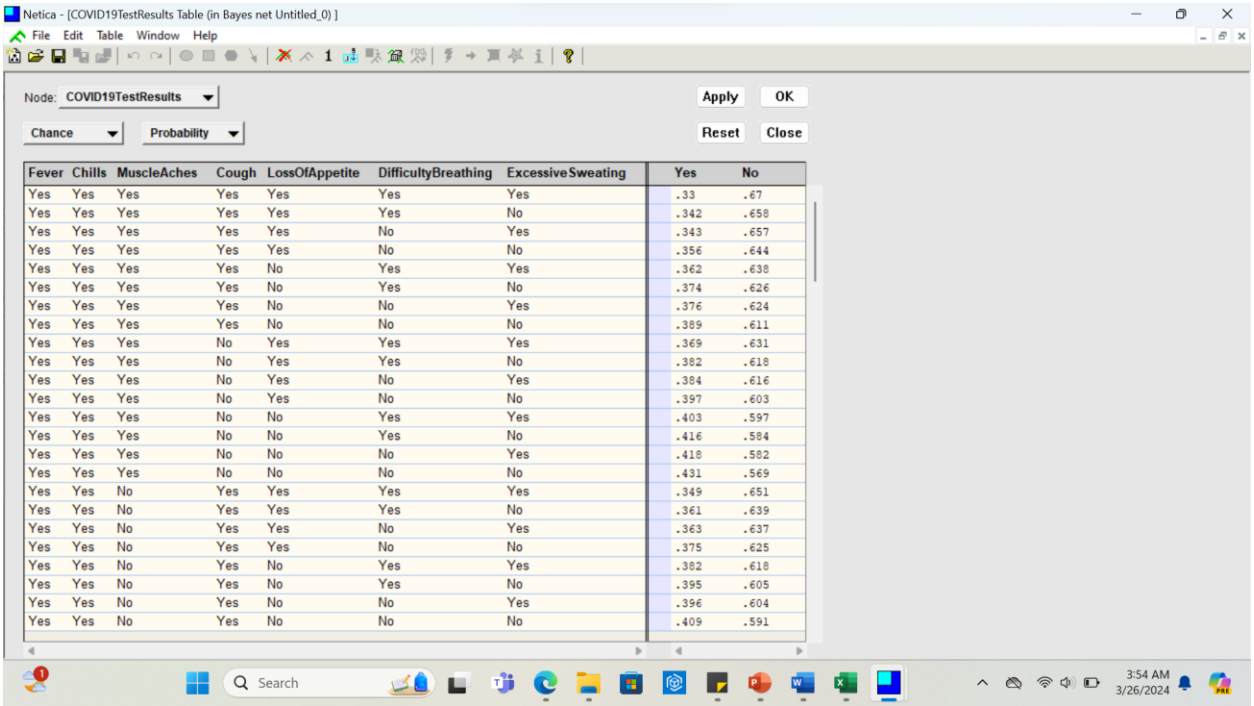
McFadden R2 Value	0.341	0	0.051	0.496	0.393	0.485	0.648	0.57	0.627	0.403	0.441	0.548	0.57	0.614	0.517	0.504	0.608	0.677	0.54	0.571	0.641
	COVID - 19 Test Results	**Shivering**	**Fatigue**	**Loss of Taste**	**Fever**	**Headaches**	**Loss of Smell**	**Chills**	**Muscle Aches**	**Diarrhea**	**Cough**	**Shortness of Breath**	**Runny Nose**	**Sore Throat**	**Loss of Balance**	**Vomiting**	**Joint Pain**	**Loss of Appetite**	**Wheezing**	**Difficulty Breathing**	**Excessive Sweating**
Intercept	0.028	0.006	0.118	0.001	0.022	0.042	0.009	0.015	0.005	0.009	0.048	-0.001	0.013	0.009	0.005	0.005	0.002	0.005	0.005	0.006	0.006
Age																					
Female																					
Shivering			0.812																		
Fatigue				0.223	0.316	0.591		0.349	0.254	0.149	0.242	0.115	0.293	0.121							0.109
Loss of Taste					0.055		0.659											0.257	0.065		
Fever	0.105					0.110	0.092	0.365													0.146
Headaches									0.071		0.339		0.110	0.150							
Loss of Smell												0.056			0.077			0.11			
Chills	0.116								0.323	0.165	0.256						0.084				
Muscle Aches	0.085									0.090		0.154				0.14	0.205	0.055			
Diarrhea																0.171		0.098			
Cough	0.174											0.064	0.257					0.099		0.108	
Shortness of Breath															0.194			0.126	0.14	0.265	
Runny Nose														0.374							
Sore Throat																					
Loss of Balance																	0.069	0.207			
Vomiting																					
Joint Pain																					0.192
Loss of Appetite	0.142																			0.236	
Wheezing																					
Difficulty Breathing	0.061																				0.241
Excessive Sweating	0.054																				

d. Estimate the parameters of the network from repeated LASSO regressions. Report the joint probability of COVID-19 positive test results, if we do not know which symptoms were present.

Answer:

Image 3: The Parameters of the Netica Network were estimated from repeated LASSO Regressions and the Joint Probability of COVID – 19 Test Results was calculated, even if the symptoms occurrence was not known, as shown below:





Fever	Chills	MuscleAches	Cough	LossOfAppetite	DifficultyBreathing	ExcessiveSweating	Yes	No
Yes	Yes	Yes	Yes	Yes	Yes	Yes	.33	.67
Yes	Yes	Yes	Yes	Yes	Yes	No	.342	.658
Yes	Yes	Yes	Yes	Yes	No	Yes	.343	.657
Yes	Yes	Yes	Yes	Yes	No	No	.356	.644
Yes	Yes	Yes	Yes	No	Yes	Yes	.362	.638
Yes	Yes	Yes	Yes	No	Yes	No	.374	.626
Yes	Yes	Yes	Yes	No	No	Yes	.376	.624
Yes	Yes	Yes	Yes	No	No	No	.389	.611
Yes	Yes	Yes	No	Yes	Yes	Yes	.369	.631
Yes	Yes	Yes	No	Yes	Yes	No	.382	.618
Yes	Yes	Yes	No	Yes	No	Yes	.384	.616
Yes	Yes	Yes	No	Yes	No	No	.397	.603
Yes	Yes	Yes	No	No	Yes	Yes	.403	.597
Yes	Yes	Yes	No	No	Yes	No	.416	.584
Yes	Yes	Yes	No	No	No	Yes	.418	.582
Yes	Yes	Yes	No	No	No	No	.431	.569
Yes	Yes	No	Yes	Yes	Yes	Yes	.349	.651
Yes	Yes	No	Yes	Yes	Yes	No	.361	.639
Yes	Yes	No	Yes	Yes	No	Yes	.363	.637
Yes	Yes	No	Yes	Yes	No	No	.375	.625
Yes	Yes	No	Yes	No	Yes	Yes	.382	.618
Yes	Yes	No	Yes	No	Yes	No	.395	.605
Yes	Yes	No	Yes	No	No	Yes	.396	.604
Yes	Yes	No	Yes	No	No	No	.409	.591

e. What are parents in the Markov blanket of Fever?

- Use regressions to identify these parents in Markov Blanket of Fever.

- Use the network to read parents in Markov Blanket of Fever.

Answer:

Image 4: Parents in the Markov Blanket of Fever using Regressions:

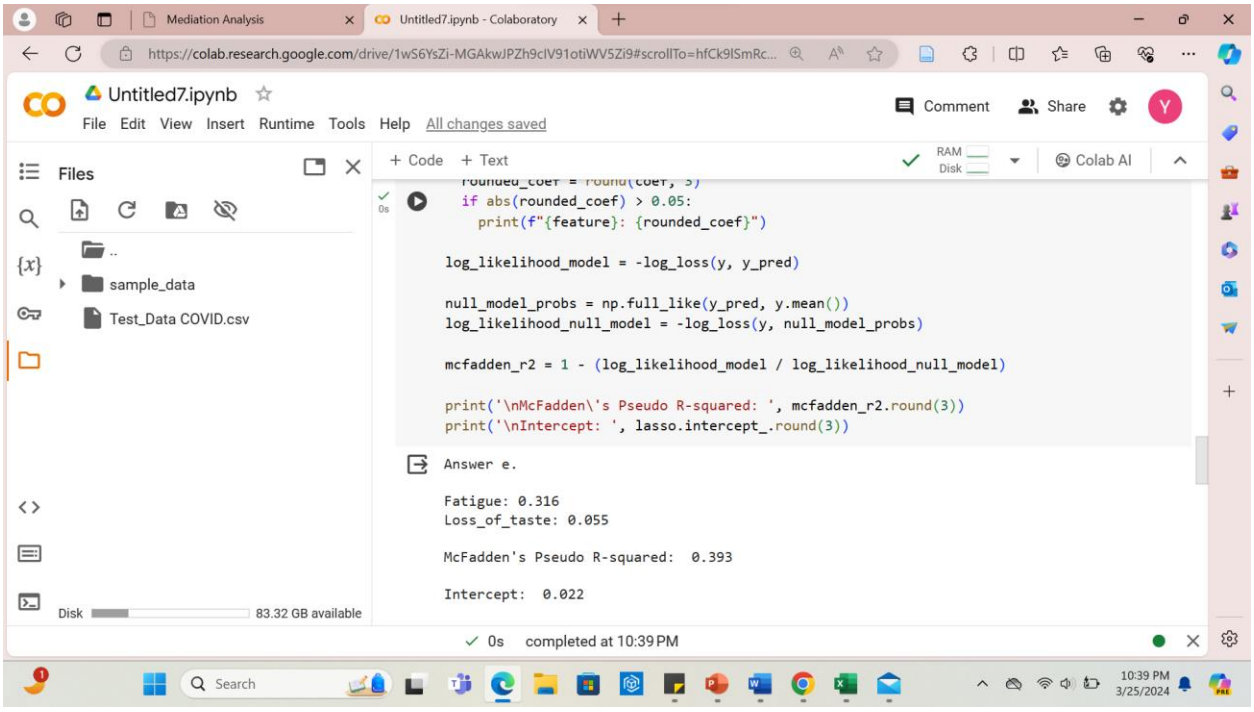
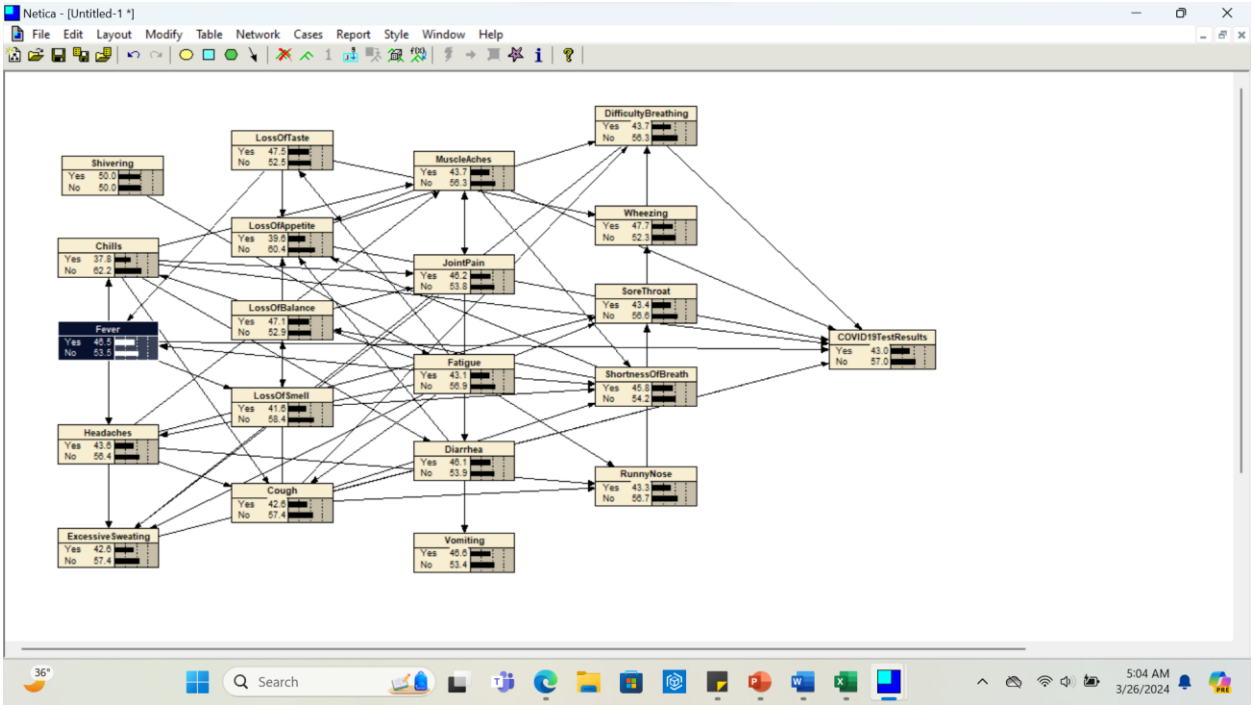


Image 5: Parents in the Markov Blanket of Fever using Netica Network:

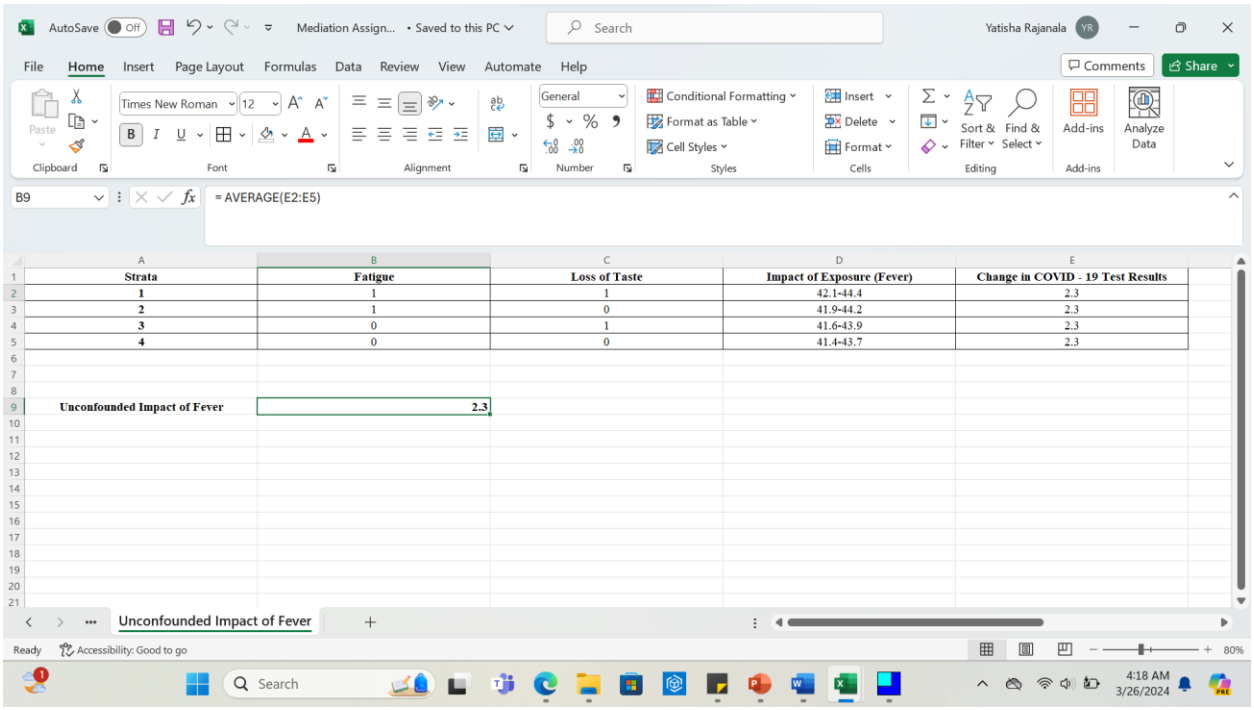


f. What is the un-confounded effect of fever on probability of positive COVID-19 diagnosis?

- Use inverse propensity weights to remove confounding.
- Switch the distribution of direct predictors of Fever so that patients with and without Fever have the same distribution of direct predictors.

Answer:

Image 6: The Unconfounded Impact of Fever on Probability of Positive COVID – 19 Diagnosis:



	A	B	C	D	E
1	Strata	Fatigue	Loss of Taste	Impact of Exposure (Fever)	Change in COVID - 19 Test Results
2	1	1	1	42.1-44.4	2.3
3	2	1	0	41.9-44.2	2.3
4	3	0	1	41.6-43.9	2.3
5	4	0	0	41.4-43.7	2.3
9	Unconfounded Impact of Fever	2.3			

Therefore, the Unconfounded Impact of Fever on Probability of Positive COVID – 19 Diagnosis is **2.3**.

g. What are the parents in Markov blanket of Chills?

- Use Network to identify the parents in Markov blanket of Chills.
- Use regressions to identify parents in Markov blanket of Chills.

Answer:

Image 7: Parents in the Markov Blanket of Chills using Netica Network:

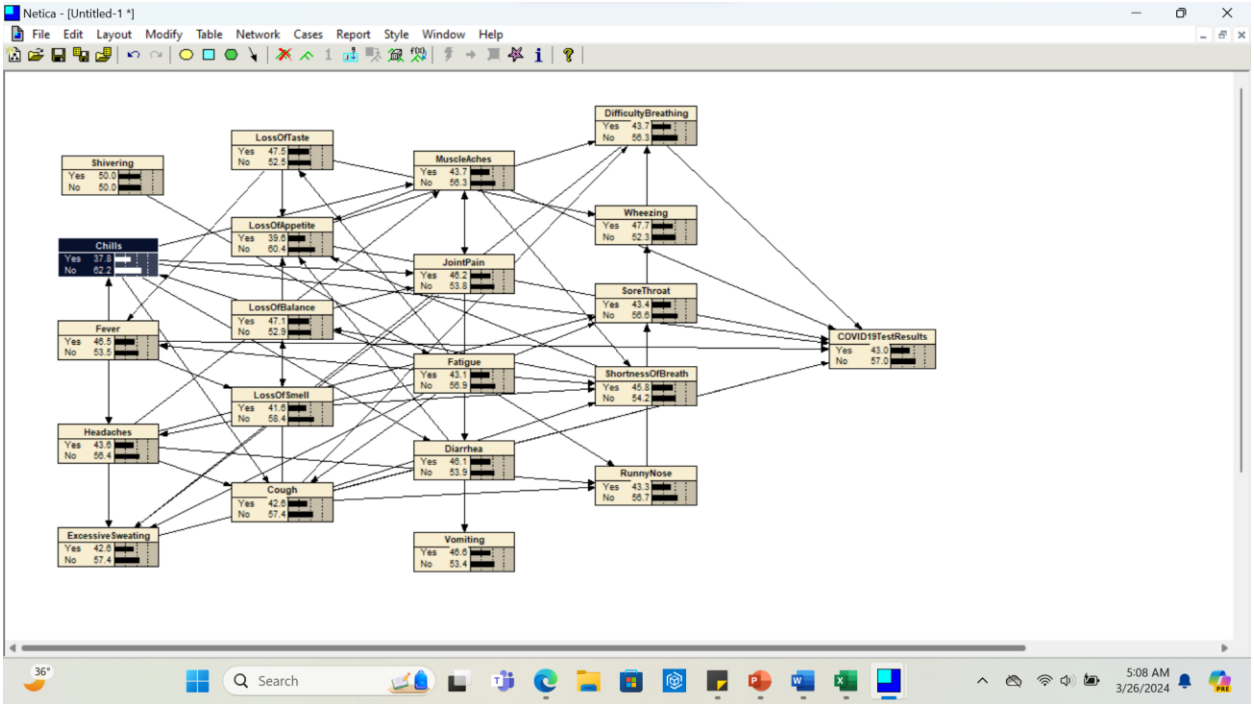
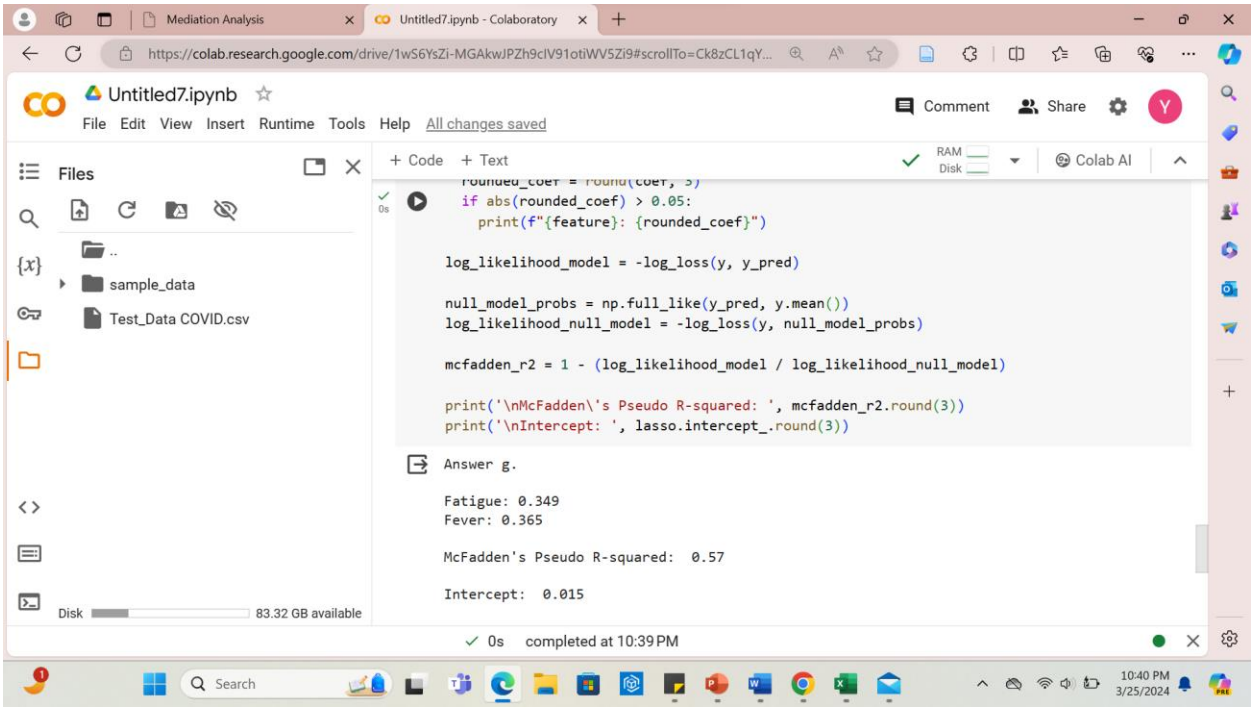


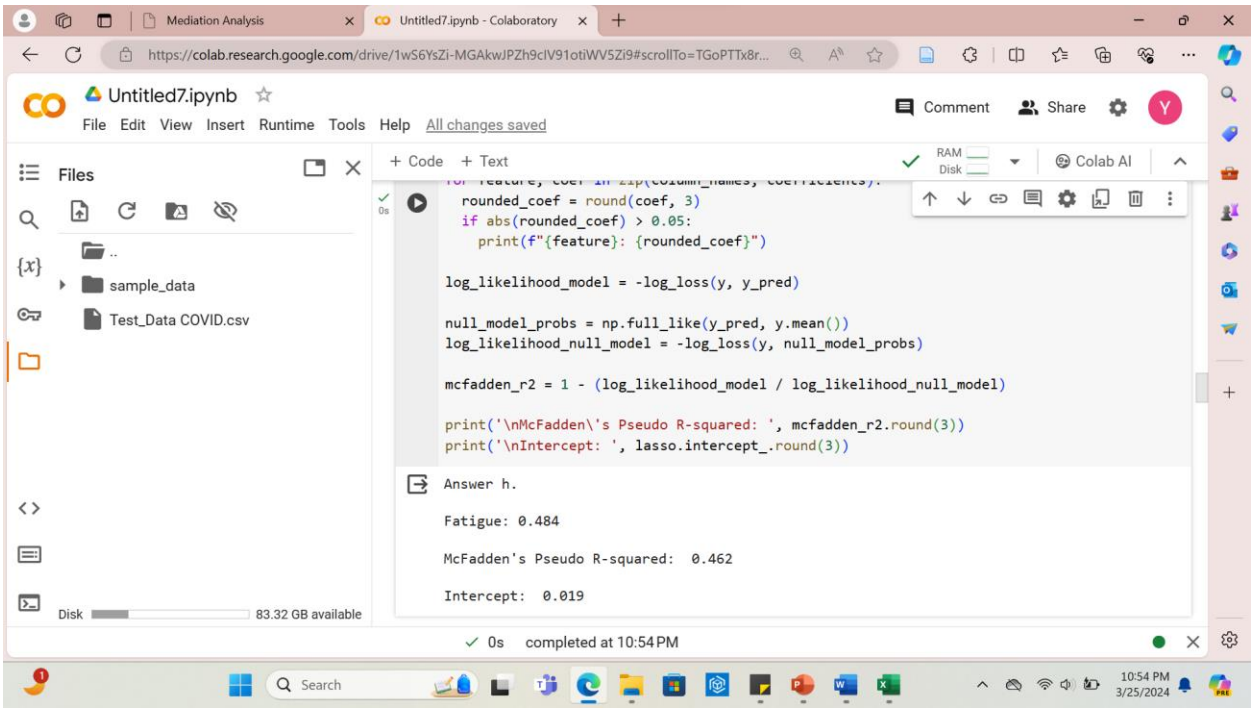
Image 8: Parents in the Markov Blanket of Chills using Regressions:



h. LASSO regress Chills on its direct predictors, not including Fever. Report intercept, coefficients, and McFadden R-square.

Answer:

Image 9: The LASSO Regression of Chills on its Direct Predictors, excluding Fever:



i. Revise the network to create a counterfactual network in which Fever is not mediated by Chills (no arc from Fever to Chills).

Answer:

Image 10: Real Netica Network:

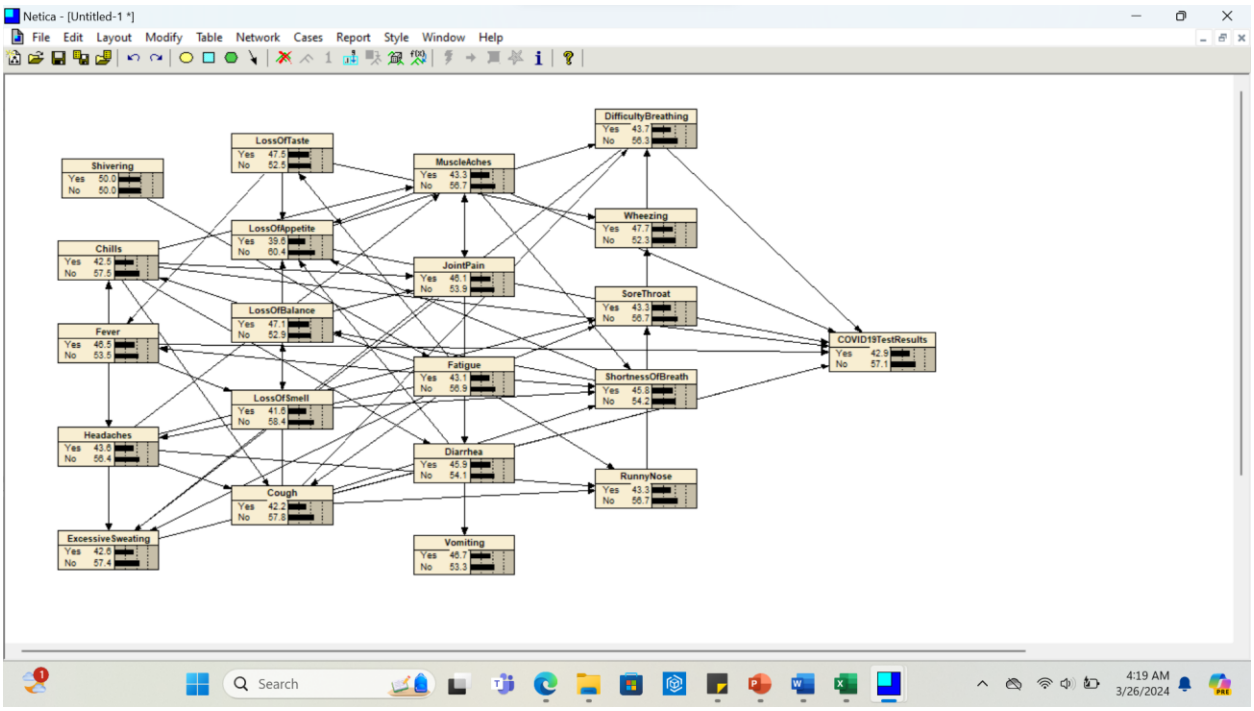
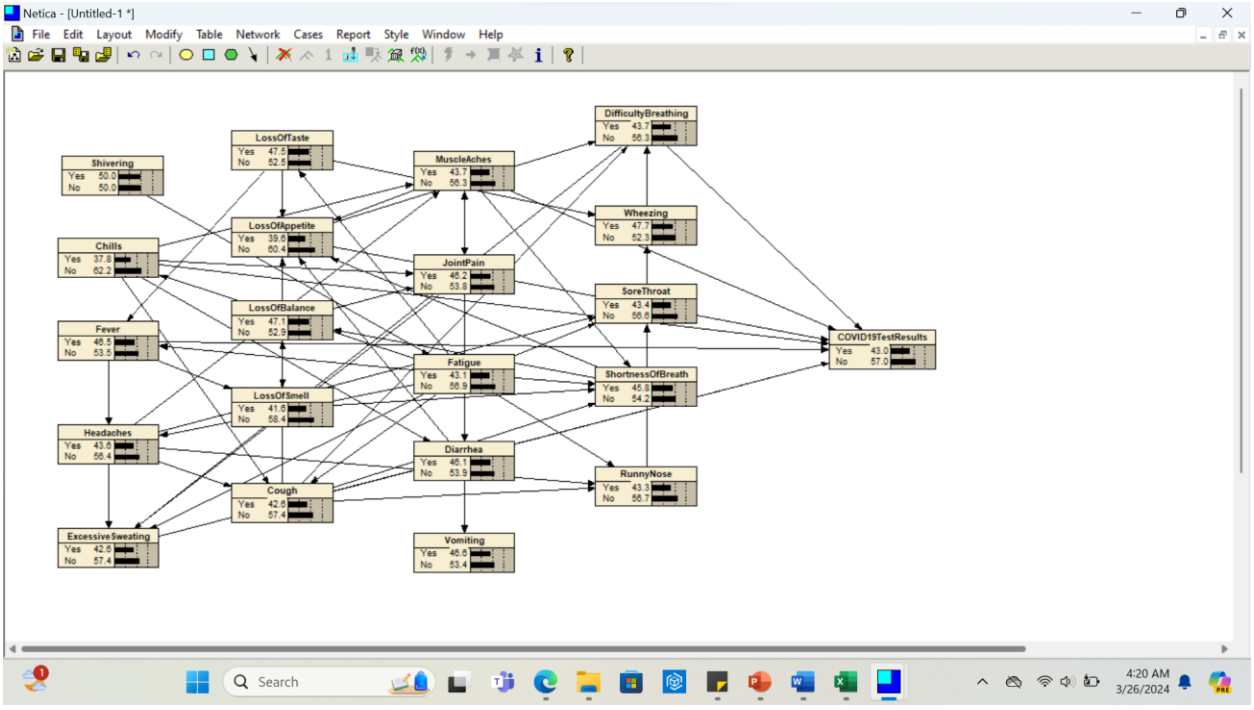


Image 11: Counterfactual Netica Network, where Fever is Not Mediated by Chills:



j. What is the mediated impact of fever on COVID-19 through Chills?

Answer:

Image 12: The Frequency of Strata can be calculated using the following:

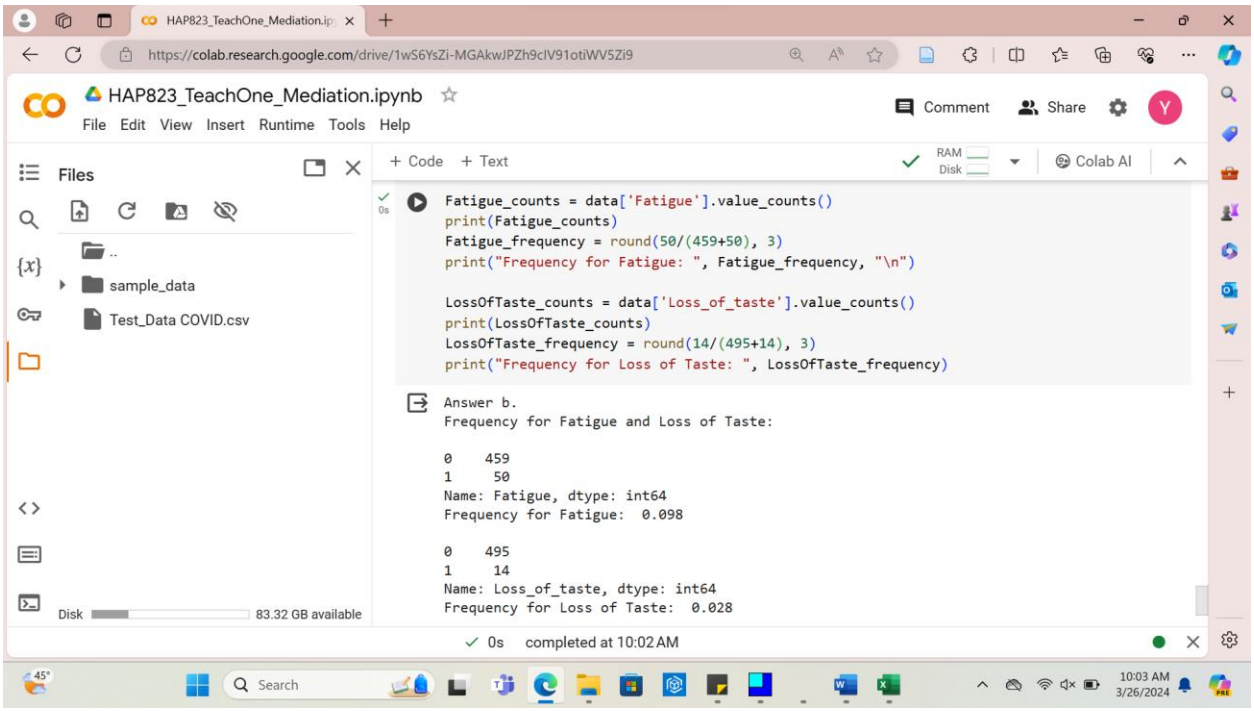
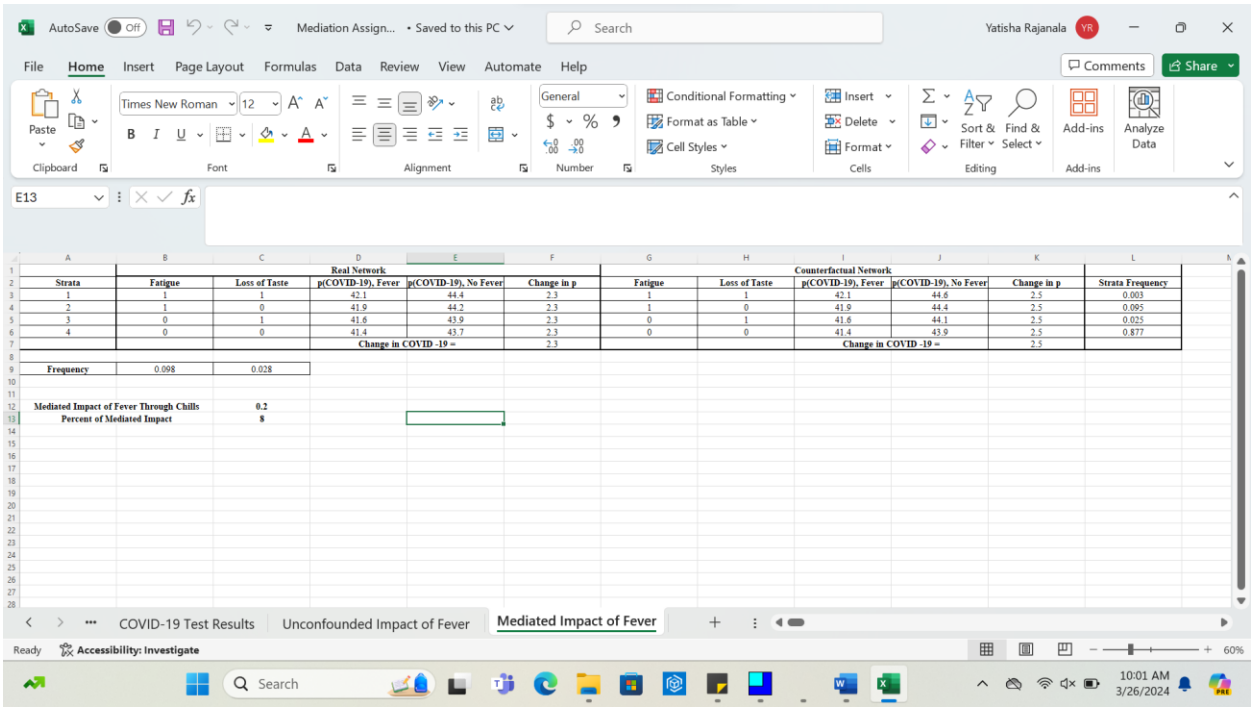


Image 13: The Mediated Impact of Fever on COVID – 19 through Chills:



			Real Network			Counterfactual Network					
Strata	Fatigue	Loss of Taste	p(COVID-19), Fever	p(COVID-19), No Fever	Change in p	Fatigue	Loss of Taste	p(COVID-19), Fever	p(COVID-19), No Fever	Change in p	Strata Frequency
1	1	1	42.1	44.4	2.3	1	1	42.1	44.6	2.5	0.003
2	1	0	41.9	44.2	2.3	1	0	41.9	44.4	2.5	0.095
3	0	1	41.6	43.9	2.3	0	1	41.6	44.1	2.5	0.025
4	0	0	41.4	43.7	2.3	0	0	41.4	43.9	2.5	0.877
			Change in COVID -19 =		2.3			Change in COVID -19 =		2.5	
Frequency	0.098	0.028									
Mediated Impact of Fever Through Chills		0.2									
Percent of Mediated Impact		8									

Therefore, the Mediated Impact of Fever on COVID – 19 through Chills is **0.2**, and the Percent of Mediated Impact is **8%**.