COMPSCI 260 - Problem Set 5, Problem 3
Due: Fri 8 Nov 2019, 5pm
Name: Kyle Mitra
NetID: km423
Statement of collaboration and resources used (put None if you worked
entirely without collaboration or resources; otherwise cite carefully): None

My solutions and comments for this problem are below.
----
a)

Index	Protein ID	Virus name	Animal host
1	AAK83356.1	Bovine coronavirus	Cattle
2	AAL57308.1	Bovine coronavirus	Cattle
3	AAL80031	Porcine hemagglutinating encephalomyelitis virus (HEV)	pigs
4	AAP41037.1	SARS coronavirus Tor2	human
5	AAR01015.1	Human coronavirus OC43	human
6	AAV49723	SARS coronavirus PC4-241	human
7	ABD75513.1	Human coronavirus HKU1	human
8	AFS88936.1	Human betacoronavirus 2c EMC/2012	human
9	ATW75480.1	Middle East respiratory syndrome-related coronavirus	Dromedary (camel)
10	AUM60024.1	Middle East respiratory syndrome-r elated coronavirus	Kuhl's pipistrelle (vesper bat)
11	CAA01637.1	Canine coronavirus	dogs
12	CAA71056.1	Human coronavirus 229E	human
13	YP 209233.1	Murine hepatitis virus strain JHM	mice

b) See code

c) Pairwise distances of the 13 proteins:

	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0051	0.1668	0.7100	0.0801	0.7100	0.3390	0.6970	0.6970	0.6900	0.7908	0.7589	0.3525
2		0.1675	0.7116	0.0801	0.7116	0.3390	0.6962	0.6962	0.6885	0.7901	0.7579	0.3517
3			0.7155	0.1910	0.7155	0.3546	0.6888	0.6888	0.6892	0.7949	0.7618	0.3583
4				0.7384	0.0127	0.7365	0.7073	0.7081	0.7019	0.8190	0.7785	0.7362
5					0.7163	0.3516	0.6962	0.6962	0.6907	0.7922	0.7618	0.3626
6						0.7358	0.7066	0.7073	0.7019	0.8204	0.7785	0.7347
7							0.7021	0.7021	0.6944	0.7908	0.7490	0.3670
8								0.0044	0.3160	0.7983	0.7569	0.7064
9									0.3167	0.7977	0.7569	0.7064
10										0.7915	0.7598	0.7006
11											0.5138	0.7842
12												0.8278

d) See code

e) _

f) _

g) _

h) _