Homework Number: 11
Name: Dylan Huynh
ECN login: huynh38
Due Date: 4/11/2023

Recipe 1: I thought I could do it with just a comma, but I needed to account for when commas are in quotes. So I used a regex to check for angle brackets facing away from each other, because you're not allowed to have them in the name of an email, and having them facing away would mean there are multiple entries.

Recipe 2: After looking through the subject lines of the emails, I noticed that most were centered around diplomas, pills, and the rest didn't really have a pattern. The ones that didn't have a matching word, I put explicitly in the regex, and caught them all with an or statement for all of them. Diploma could be with a "D" or "d" so I put an or statement for both cases, and pills, in between "P/p" and "ills", were random letters, so I put a caught them with that, and then I put a P|p just because it felt right to actually catch the word "pills".

Recipe 3: I found that the body of the 3 emails had a money based theme to them, so I took GOLD and Ruby and put them into a regex to filter them out. I had to put a body flag on the colon line to make sure the recipe knew to only look at the body.

Recipe 4: This one was a compound rule based on the MIME header. I saw that all of the junk mail had a multipart/alternative content type so I first caught that. Then I made sure that it contained the two types, plain and html. Then I check that it either used quoted-printable|8bit|7bit for its encoding type. Unfortunately, a normal email sent from Purdue Outlook shares these 4 conditions, but what's different is that it's charset is us-ascii, so I added to the rules for plain and html that their charset must either be "utf-8", utf-8, or nothing is written after the semicolon.