Alphabetic only texts/strings

For the core parts of the Hybrid Cryptography project you are expected to work with text strings that only contain alphabetic characters, i.e. no spaces, no numbers, no punctuation, just the letters of the alphabet.

You will therefore need to write a function that extracts precisely the alphabetic content of a string. Let's suppose that you have done this and that you have defined the following function

```
def extract_text(text):
    # Body of the function which takes the string
    # text and returns a string containing only the
    # alphabetic content of text. For you to define...
```

Now as an example we will define the string first_part to be the string containing the first 4000 characters of the string austen_text whose content is the book *Pride and Prejudice* by Jane Austen. (4000 characters chosen so that we have started the actual book or, at least, the preface!) To do this we simply repeat the work done at the beginning of the notebook get_online_texts.ipynb where we noted that we could find the *utf-8* encoded version of *Pride and Prejudice* by Jane Austen here:

https://www.gutenberg.org/files/1342/1342-0.txt

and where we downloaded this as a string using the function url_to_text_utf8 defined below.

So now let's get the book Pride and Prejudice (as a string) once again.

```
In [6]: austen_text = url_to_text_utf8("https://www.gutenberg.org/files/1342/1342-0.txt")
```

And continue by defining first_part.

```
In [7]: first_part = austen_text[:4000]
```

Now let's have a look at this string. Notice all the space and other junk...

```
In [8]: first_part
```

Out[8]:

'The Project Gutenberg eBook of Pride and prejudice, by Jane Austen\r\n\r\nThis eBook is for the use of anyone anywhere in the United States and\r\nmost other parts of the world at no cost and with almost no restrictions\r\nwhatsoever. You may copy it, give it away or re-use it under the terms\r\nof the Projec t Gutenberg License included with this eBook or online at\r\nwww.gutenberg.org. If you are not located in the United States, you\r\nwill have to check the laws of the country where you are located before\r \r\nProduced by: Chuck Greif and the Online Distributed Proofreading Team at\r\n http://ww w.pgdp.net (This file was produced from images\r\n available at The Internet Archive)\r\n\r \n*** START OF THE PROJECT GUTENBERG EBOOK PRIDE AND PREJUDICE ***\r\n\r\n\r\n\r\n\r\n\r\n [Illustration:\r\n\r\n GEORGE ALLEN\r\n PUBLI 156 CHARING CROSS ROAD\r\n SHER\r\n\r\n LONDON\r $\n\r\n$ RUSKIN HOUSE\r\n]\r\n\r\n [Illustration:\r\n\r\n Reading Jane's Letters. Chap 34. \r\n PRIDE.\r\n $]\r\n\r\n\r\n\r\n$ and\r PREJUDICE\r\n\r\n bv\r\n with a Preface bv\r\n Jane Austen.\r\n\r\n George S $aintsbury\r\n$ and\r\n Illustrations by \r Hugh Thomson $\r\$ [Illustration: 1894]\r\n\r\n Ruskin 156. Charing $\r\$ House. Cross Road.\r\n\r\n London\r\n George Allen. $\r\n\r\n\r\n\r\n\r\n$ CHISWICK PRESS: -- C HARLES WHITTINGHAM AND CO.\r\n TOOKS COURT, CHANCERY LANE, LONDON.\r\n\r\n\r\n\r\n [Illustration:\r\n\r\n To J. Comyns Carr\r\n in acknowle dgment of all $I\r\n$ owe to his friendship and \n advice, thes e illustrations are \r n gratefully inscribed $\r\n\$]\r\n\r\n\r\n\r\n\r\nPREFACE.\r\n\r\n[Illustratio Hugh Thomson \r\n n]\r\n\r\n_Walt Whitman has somewhere a fine and just distinction between "loving\r\nby allowance" and "loving with personal love." This distinction applies\r\nto books as well as to men and women; and in the case of the not very \n numerous authors who are the objects of the personal affection, it \n nbr ings a curious consequence with it. There is much more difference as\r\nto their best work than in the case of those others who are loved "by\r\nallowance" by convention, and because it is felt to be the ri ght and\r\nproper thing to love them. And in the sect--fairly large and yet\r\nunusually choice--of Aus tenians or Janites, there would probably be\r\nfound partisans of the claim to primacy of almost every one of the\r\nnovels. To some the delightful freshness and humour of_ Northanger\r\nAbbey, _its complet eness, finish, and entrain, obscure the undoubted\r\ncritical facts that its scale is small, and its scheme, after all, that\r\nof burlesque or parody, a kind in which the first rank is reached with\r\ndi $fficulty._ \ Persuasion, \ _relatively \ faint \ in \ tone, \ and \ not \ enthralling \verb|\|r| nin \ interest, \ has \ devotees \ who$ exalt above all the others its exquisite\r\ndelicacy and keeping. The catastrophe of Mansfield Park _i s admittedly\r\ntheatrical, the hero and heroine are insipid, and the author has almost\r\nwickedly des troyed all romantic interest by expressly admitting that\r\nEdmund only took Fanny because

Extract the alphabetic content of the string first_part

To extract the alphabetic content (i.e. alphabetic characters) from the string first_part we pass it to the function extract_text. Using the following line of code we assign the variable alphabetic_first_part to the resulting string.

```
alphabetic_first_part = extract_text(first_part)
```

And now, when we inspect the string returned by extract_text by, for example, running a cell containing only the line...

```
alphabetic_first_part
```

we see — provided extract_text is correctly defined — that this string is as follows:

 $\verb|'TheProjectGutenbergeBookofPrideandprejudicebyJaneAustenThiseBookisfortheuseofanyoneanywhereing the project of the project$ $\verb|ntheUnitedStates| and \verb|mostotherparts| of the \verb|worldat| no cost and \verb|withalmostnorestrictions| what so ever \verb|Youma| and the total cost and$ ycopyitqiveitawayorreuseitunderthetermsoftheProjectGutenberqLicenseincludedwiththiseBookoronl $in eatwww guten bergorg {\tt Ifyouare not located} in the {\tt United Statesyouwill have to check the laws of the country weights to the {\tt United Statesyouwill have to check the laws of the {\tt Country Weights} in {\tt Country Weight$ $here you are located before using thise {\tt BookTitlePride} and {\tt prejudiceAuthorJaneAustenReleaseDateNovembe}$ re Book Most recently updated March Language English Produced by Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced by Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced by Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced by Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced by Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced By Chuck Greif and the Online Distributed Proofing Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English Produced By Chuck Greif and Control of the March Language English By Chuck Greif Andrew Control of the March Language English By Chuck Greif and Control of the March Language English By Chuck Greif and Control of the March Language English By Chuck Greif Andrew Control of the March Language English By Chuck Greif Andrew Control of the March Language English By Chuck Greif Andrew Control of the March Language E $reading {\tt TeamathttpwwwpgdpnetThisfile} was {\tt producedfromimages} available at {\tt TheInternetArchiveSTARTOFT}$ ${\tt HEPROJECTGUTENBERGEBOOKPRIDEANDPREJUDICEIllustrationGEORGEALLENPUBLISHERCHARINGCROSSROADLONDO}$ NRUSKINHOUSEIllustrationReadingJanesLettersChapPRIDEandPREJUDICEbyJaneAustenwithaPrefacebyGeo ${\tt rgeSaintsburyandIllustrationsbyHughThomsonIllustrationRuskinCharingHouseCrossRoadLondonGeorgeng} \\$ AllenCHISWICKPRESSCHARLESWHITTINGHAMANDCOTOOKSCOURTCHANCERYLANELONDONIllustrationToJComynsCar rinacknowledgmentofallIowetohisfriendshipandadvicetheseillustrationsaregratefullvinscribedHug $\verb|hThomsonPREFACEIllustrationWaltWhitman| has somewhere a fine and just distinction between loving by allow a constraint of the property of$ $ance and loving with personal love {\tt This distinction} applies to books as {\tt well as to men and women and in the case of the total content of the total$ the not very numerous authors who are the objects of the personal affection it brings a curious consequence with the notion of the personal affection in the personal affect $\verb|hitThere is much more difference as to the \verb|irbestwork| than in the case of those others who are loved by allowance of the loved by allowance o$ byconventionandbecauseitisfelttobetherightandproperthingtolovethemAndinthesectfairlylargeandy $et unu sually choice of {\tt Austeniansor Janites the rewould probably be found part is an soft heclaim to primacy of the control of the cont$ $almost every one of the novels {\tt To} somethed elightful freshness and humour of {\tt NorthangerAbbeyits} completenes$ $s finish and entrain obscure the undoubted critical facts that its scale is {\tt small} and its {\tt scheme} after all that of the {\tt scale} is {\tt small} and {\tt scheme} after all that of {\tt scale} is {\tt scale} is {\tt small} and {\tt scale} is {$ $fburles que or paro dyak ind in which the first rank is reached with difficulty {\tt Persuasion} relatively {\tt faintint}$ oneandnotenthrallingininteresthasdevoteeswhoexaltabovealltheothersitsexquisitedelicacyandkeep ing The catastrophe of Mansfield Parkis admitted ly the atrical thehero and hero in ear einsipid and the author of the property of the prope $has almost wicked ly destroyed all romantic interest by expressly admitting that {\tt Edmundonly took} Fanny because the {\tt Edmundonly took} fanny because {\tt Edmundonly took} fanny {\tt$ se

The Point

You can convert all of the string austen_text in this way to obtain a purely alphabetic string then use it for encryption and decryption etc. (and/or save it in a file for use later on). You can (and should) of course follow the same process for any of the text (.txt) files on the Hybrid Cryptography GitHub page, or for any other text files that you choose to use.

Note on cases

You can design your encryption and decryption functions so that they preserve case. However you can also decide to work only in upper case or lower case (although preservation of case would be a nice feature to have). To do this you would simply have to convert your string to lower or upper case.

For example executing the line

```
alphabetic_first_part.lower()
returns the string:
```

'theprojectgutenbergebookofprideandprejudicebyjaneaustenthisebookisfortheuseofanyoneanywherei $\verb|ntheunited| states| and \verb|mostotherparts| of the \verb|worldatnocostand| with almost no restrictions what so every our manufacture of the terms of t$ ycopyitgiveitawayorreuseitunderthetermsoftheprojectgutenberglicenseincludedwiththisebookoronl in eatwww guten bergorg if you are not located in the united states you will have to check the laws of the country will be a constant of the constant of the country will be a constant of the chere you are located before using this ebook title pride and prejudice authorjane austen released at enovember and the pride arebook most recently updated march language english produced by chuck greif and the online distributed proof the contraction of the contraction $reading team at \verb|http| wwwpgdpnetth is filewas produced from images a vailable at the internet archive start of the produced from the p$ he project gut en bergebook pride and prejudice illustration george all en publisher charing cross road londomore in the project gut en pronruskinhouseillustrationreadingjanesletterschapprideandprejudicebyjaneaustenwithaprefacebygeo rgesaintsburyandillustrationsbyhughthomsonillustrationruskincharinghousecrossroadlondongeorge all enchis wick presscharles whitting ham and cotooks court chancery lanel on donillustration to jcomyns carries and the contraction of the contrinacknowledgmentofalliowetohisfriendshipandadvicetheseillustrationsaregratefullyinscribedhug hthomsonprefaceillustrationwaltwhitmanhassomewhereafineandjustdistinctionbetweenlovingbvallow ance and loving with personal love this distinction applies to books as well as to men and women and in the case of the control of the contthe not very numerous authors who are the objects of the personal affection it brings a curious consequence with the contraction of the contracthitthereismuchmoredifferenceastotheirbestworkthaninthecaseofthoseotherswhoarelovedbyallowance $by convention and because \verb|itisfe|| tho bethe right and proper thing to love the mand in the sect fairly large and years of the love them and the sect fairly large and years of the love them are the love them and the love them are the love the love them are the love the love them are the love them are the love them are the love the love them are the love them are the love the love the love them are the love the love the love them are the love the$ et unu sually choice of a ustenian sor janites the rewould probably befound part is an softheclaim to primacy of the contraction of the contractalmostevery one of the novel sto somethed elightful freshness and humour of northanger abbey its completenessfinish and entrain obscure the undoubted critical facts that its scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scheme after all that of the scale is small and its scale is small and its scheme after all the scale is small and its scheme after all that of the scale is small and its scale isfburles que or parody a kindin which the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with difficulty persuasion relatively faint in the first rank is reached with the first rank is reached with difficulty persuasion relatively faint rank is reached with the first rank is reached wione and notenth rall ling in interest has devotees who exalt above all the other sits exquisited elicacy and keep the content of the conten $ing the catastrophe of {\tt mansfield} park is admittedly the {\tt atrical} the hero and hero in {\tt eare} in {\tt sipid} and the {\tt author} is {\tt otherwise} and {\tt otherwise} and$ has almost wicked ly destroyed all romantic interest by expressly admitting that edmund only took fanny because the substitution of the contraction of the contract

```
whereas executing the line
alphabetic_first_part.upper()
returns the string
```

'THEPROJECTGUTENBERGEBOOKOFPRIDEANDPREJUDICEBYJANEAUSTENTHISEBOOKISFORTHEUSEOFANYONEANYWHEREI NTHEUNITEDSTATESANDMOSTOTHERPARTSOFTHEWORLDATNOCOSTANDWITHALMOSTNORESTRICTIONSWHATSOEVERYOUMA YCOPYITGIVEITAWAYORREUSEITUNDERTHETERMSOFTHEPROJECTGUTENBERGLICENSEINCLUDEDWITHTHISEBOOKORONL INEATWWWGUTENBERGORGIFYOUARENOTLOCATEDINTHEUNITEDSTATESYOUWILLHAVETOCHECKTHELAWSOFTHECOUNTRYW HEREYOUARELOCATEDBEFOREUSINGTHISEBOOKTITLEPRIDEANDPREJUDICEAUTHORJANEAUSTENRELEASEDATENOVEMBE ${\tt REBOOKMOSTRECENTLYUPDATEDMARCHLANGUAGEENGLISH PRODUCED BY CHUCKGREIF AND THE ONLINED IS TRIBUTED PROOFFICE AND THE OWNER OF THE OWNER OWNER$ READINGTEAMATHTTPWWWPGDPNETTHISFILEWASPRODUCEDFROMIMAGESAVAILABLEATTHEINTERNETARCHIVESTARTOFT ${\tt HEPROJECTGUTENBERGEBOOKPRIDEANDPREJUDICEILLUSTRATIONGEORGEALLENPUBLISHERCHARINGCROSSROADLONDO$ ${\tt NRUSKINHOUSEILLUSTRATIONREADINGJANESLETTERSCHAPPRIDEANDPREJUDICEBYJANEAUSTENWITHAPREFACEBYGEO$ RGESAINTSBURYANDILLUSTRATIONSBYHUGHTHOMSONILLUSTRATIONRUSKINCHARINGHOUSECROSSROADLONDONGEORGE ${\tt ALLENCHISWICKPRESSCHARLESWHITTINGHAMANDCOTOOKSCOURTCHANCERY LANELONDONIL LUSTRATION TO JCOMYNSCAR \\$ RINACKNOWLEDGMENTOFALLIOWETOHISFRIENDSHIPANDADVICETHESEILLUSTRATIONSAREGRATEFULLYINSCRIBEDHUG HTHOMSONPREFACEILLUSTRATIONWALTWHITMANHASSOMEWHEREAFINEANDJUSTDISTINCTIONBETWEENLOVINGBYALLOW ${\tt ANCEANDLOVINGWITHPERSONALLOVETHISDISTINCTION APPLIESTOBOOKS AS WELLASTOMEN AND WOMEN AND INTHE CASE OF THE STATE OF T$ THENOTVERYNUMEROUSAUTHORSWHOARETHEOBJECTSOFTHEPERSONALAFFECTIONITBRINGSACURIOUSCONSEQUENCEWIT HITTHEREISMUCHMOREDIFFERENCEASTOTHEIRBESTWORKTHANINTHECASEOFTHOSEOTHERSWHOARELOVEDBYALLOWANCE ${\tt BYCONVENTIONANDBECAUSEITISFELTTOBETHERIGHTANDPROPERTHINGTOLOVETHEMANDINTHESECTFAIRLY LARGE AND YOUR CONTROL OF THE SECOND O$ ${\tt ETUNUSUALLYCHOICEOFAUSTENIANSORJANITESTHEREWOULDPROBABLYBEFOUNDPARTISANSOFTHECLAIMTOPRIMACYOFICAL CONTROL OF CONTROL$ ALMOSTEVERYONEOFTHENOVELSTOSOMETHEDELIGHTFULFRESHNESSANDHUMOUROFNORTHANGERABBEYTTSCOMPLETENES ${\tt SFINISHANDENTRAINOBSCURETHEUN DOUBTED CRITICAL FACTS THAT ITS SCALE IS SMALL AND ITS SCHEME AFTER ALL THAT OF STANDARD CONTRACTOR OF$ FBURLESQUEORPARODYAKINDINWHICHTHEFIRSTRANKISREACHEDWITHDIFFICULTYPERSUASIONRELATIVELYFAINTINT ONEANDNOTENTHRALLINGININTERESTHASDEVOTEESWHOEXALTABOVEALLTHEOTHERSITSEXOUISITEDELICACYANDKEEP $\textbf{INGTHECATASTROPHEOFMANSFIELDPARKISADMITTEDLYTHEATRICALTHEHEROANDHEROINEAREINSIPIDANDTHEAUTHOR$ HASALMOSTWICKEDLYDESTROYEDALLROMANTICINTERESTBYEXPRESSLYADMITTINGTHATEDMUNDONLYTOOKFANNYBECAU SE