Reading file formats

To read a text file (.txt) save it in the directory of the notebook, choose a variable, equal to open('nameofyourfiel.txt'), then as follows: nameofvariable.read() use the function to read the text (see the notebook for more details).

```
In [1]: diotima = open('diotima.txt')

In [2]: diotima.read()

Out[2]: ' I said, "O thou stranger woman, thou sayest well; but, assuming Love\nto be such as you say, what is the use of him to me
        n?" "That,\nSocrates," she replied, "I will attempt to unfold: of his nature and\nbirth I have already spoken; and you acknow
        ledge that love is of the\nbeautiful. But some one will say: Of the beautiful in what, Socrates\nand Diotima?-or rather let m
        e put the question more dearly, and ask:\nWhen a man loves the beautiful, what does he desire?" I answered her\n"That the bea
        utiful may be his." "Still," she said, "the answer\nsuggests a further question: What is given by the possession of\nbeauty?"
        "To what you have asked," I replied, "I have no answer\nready." "Then," she said, "Let me put the word \'good\' in the place
        \nof the beautiful, and repeat the question once more: If he who loves\ngood, what is it then that he loves? "The possession
        of the good," I\nsaid. "And what does he gain who possesses the good?" "Happiness," I\nreplied; "there is less difficulty in
        answering that question." "Yes,"\nshe said, "the happy are made happy by the acquisition of good things.\nNor is there any ne
        ed to ask why a man desires happiness; the answer\nis already final." "You are right." I said. "And is this wish and this\nde
        sire common to all? and do all men always desire their own good,\nor only some men?-what say you?" "All men," I replied; "the
        desire\nis common to all." "Why, then," she rejoined, "are not all men,\nSocrates, said to love, but only some them? whereas
        you say that all\nmen are always loving the same things." "I myself wonder," I said,-why\nthis is." "There is nothing to wond
        er at," she replied; "the reason is\nthat one part of love is separated off and receives the name of the\nwhole, but the othe
        r parts have other names." "Give an\nillustration," I said. She answered me as follows: "There is poetry,\nwhich, as you kno
        w, is complex; and manifold. All creation or\npassage of non-being into being is poetry or making, and the processes\nof all
        art are creative; and the masters of arts are all poets or\nmakers." "Very true." "Still," she said, "you know that they are
        not\ncalled poets, but have other names; only that portion of the art which\nis separated off from the rest, and is concerned
        with music and metre,\nis termed poetry, and they who possess poetry in this sense of the\nword are called poets." "Very tru
```
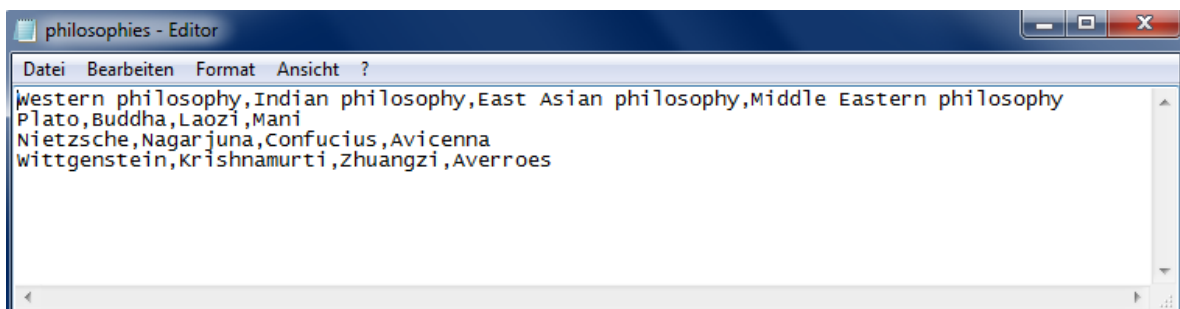
To read text files of other formats, copy the following in a cell and replace the name of your file, which you must save there where you save the notebooks. You can replace "content" with another word, but to show the text, you must repeat this name in another cell and, of course, run the cells.

```
In [ ]: import tika
        tika.initVM()
        from tika import parser
        parsed = parser.from_file('yourfile.pdf')
        print(parsed["metadata"])
        content=parsed["content"]

In [ ]: content
```

CSV-File

You have to get or make a CSV-File. Open up any text editor and enter your values all separated by commas. Save it as a CSV-File in the directory where your notebooks are saved.

```
philosophies - Editor
Datei  Bearbeiten  Format  Ansicht  ?
Western philosophy,Indian philosophy,East Asian philosophy,Middle Eastern philosophy
Plato,Buddha,Laozi,Mani
Nietzsche,Nagarjuna,Confucius,Avicenna
Wittgenstein,Krishnamurti,Zhuangzi,Averroes
```

Use the function and run the cell:

```python
name_DataFrame = pd.read_csv('nameOfYourFile.csv')
name_DataFrame
```

```
In [3]: philosophers = pd.read_csv('philosophies.csv')
```

```
In [4]: philosophers
```

Out[4]:

	Western Philosophy	Indian Philosophy	East Asian philosophy	Middle Eastern philosophy
0	Plato	Buddha	Laozi	Mani
1	Nietzsche	Nagarjuna	Confucius	Avicenna
2	Wittgenstein	Krishnamurti	Zhuangzi	Averroes

JSON

- Here is an example:

```
In [4]: # Example of what a JSON (JavaScript Object Notation) looks like:
        js = """
        {   "Branch": ["Metaphysics","Aesthetics","Ethics","Logic"],
            "Philosophers": ["Metaphysicians","Aestheticians","Ethicists","Logicians"],
            "Example": ["David Lewis","Immanuel Kant","Aristotle","William of Ockham"],
            "School": ["Analytic","Continental","Peripatetic","Scholasticism"]
        }
        """
```

```
In [5]: import json
```

```
In [6]: data = json.loads(js)
```

```
In [7]: data
```

```
Out[7]: {'Branch': ['Metaphysics', 'Aesthetics', 'Ethics', 'Logic'],
         'Philosophers': ['Metaphysicians', 'Aestheticians', 'Ethicists', 'Logicians'],
         'Example': ['David Lewis', 'Immanuel Kant', 'Aristotle', 'William of Ockham'],
         'School': ['Analytic', 'Continental', 'Peripatetic', 'Scholasticism']}
```

```
In [8]: #Convert back to JSON
        json.dumps(data)
```

```
Out[8]: '{"Branch": ["Metaphysics", "Aesthetics", "Ethics", "Logic"], "Philosophers": ["Metaphysicians", "Aestheticians", "Ethicists",
        "Logicians"], "Example": ["David Lewis", "Immanuel Kant", "Aristotle", "William of Ockham"], "School": ["Analytic", "Continenta
        l", "Peripatetic", "Scholasticism"]}'
```

- Do not forget to import the json module:

```
import json
```

- With the following you can load json data:

```
choose_name = json.loads(name_of_your_json_obj)
```

- With the following you can open JSON data after loading with a DataFrame:

```
choose_another_name = DataFrame(choose_name)
```

```
In [9]: dframe = DataFrame(data)
```

```
In [10]: dframe
```

Out[10]:

	Branch	Philosophers	Example	School
0	Metaphysics	Metaphysicians	David Lewis	Analytic
1	Aesthetics	Aestheticians	Immanuel Kant	Continental
2	Ethics	Ethicists	Aristotle	Peripatetic
3	Logic	Logicians	William of Ockham	Scholasticism

- To show the DataFrame, write again your chosen name of the variable and run the cell.