



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Proseminar “Softwareentwicklung in der Wissenschaft”

Hausarbeit zum Thema:

Jupyter Notebook

Software-Tool für „interactive data science and scientific computing“

Vorgelegt von:

Khorshid Biria

Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Informatik
Arbeitsbereich Wissenschaftliches Rechnen

Studiengang: Informatik
Matrikelnummer: 7111693

Betreuer: Dr. Hermann Lenhart

Hamburg, 2020-08-19

Inhaltsverzeichnis

Einführung	2
Jupyter Notebook in Begrifflichkeiten	3
Projekt Jupyter	4
Schema und Anwendungen in einem Beispielprogramm.....	4
Funktion und Funktionalität	8
Vorteile	9
Zusammenfassung	10
Literaturverzeichnis.....	11

Einführung

Jupyter Notebook ist eine webbasierte Client-Server-Anwendung der Non-Profit-Organisation Project Jupyter. Dieses Projekt wurde 2014 von Fernando Pérez gegründet. [wiki1]

Diese Organisation wurde dazu gegründet, um eine Open-Source-Software, die den Klienten beim Bearbeiten und Ausführen eines Fremdcodes hilft, zu erschaffen.

Verschiedene Sprachen sowie Python oder Java werden von dieser Umgebung unterstützt. So stellt uns Jupyter Notebook eine perfekte Umgebung zur Verfügung, in der man jeglichen Code übersetzen lassen kann.

Sowohl für die Datenvisualisierung als auch fürs Dokumentieren, Berechnungen Durchführen und Resultate Bekommen eignet sich diese Software in jeder Hinsicht. [ionos]

Diese Hausarbeit ist so strukturiert, dass der Leser erstmal einen ersten Überblick über das Programm Jupyter Notebook gewinnt, und eine genauere Einsicht in den wichtigsten Eigenschaften davon bekommt.

Gefolgt von einer kurzen Vorstellung des gesamten Projektes „Jupyter“ kommt die Arbeit zu dem Abschnitt „Schema und Anwendungen in einem Beispielprogramm“, in dem es vorgestellt wird, wie die Oberfläche von dem Programm genau aussieht, und was für Möglichkeiten und Anwendungen uns da zur Verfügung gestellt sind.

Schließlich wird über die Funktion und Funktionalität von Jupyter Notebook und die Vorteile davon berichtet.

Jupyter Notebook in Begrifflichkeiten

✚ Eine webbasierte Umgebung:

So eine Umgebung _auch bekannt als „Webanwendung“_ ist ein Anwendungsprogramm nach dem Client-Server-Modell. Die Datenverarbeitung findet dadurch auf einem entfernten Webserver statt, sodass die Webanwendungen nicht zwingend lokal auf dem Rechner des Benutzers installiert werden müssen. (siehe Fig. 1)

Dazu braucht der Benutzer einen Webbrowser, wodurch die Datenübertragung mit dem Webserver über das HTTP-Protokoll läuft. [wiki2]

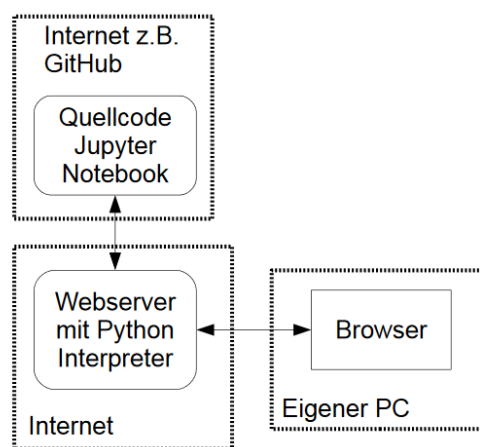


Fig. 1: Client-Server-Modell - [nwt]

✚ Eine Open-Source-Software (OSS):

Mit Open-Source-Software ist eine Software gemeint, deren Code beliebig kopiert, genutzt und verändert werden kann. Eine umfassende freie Nutzung dieser Software ist durch deren Lizenz gestattet, d.h., dass die Software zu allen Zwecken frei benutzt werden kann. [ifros]

✚ Eine wichtige Komponente vom Jupyter Notebook:

Über 50 unterschiedliche Sprachen werden vom Jupyter Notebook unterstützt. Dieses Programm kann nämlich verschiedene Kernels aufrufen, von daher können verschiedene Codes in der jeweiligen Sprache ausgeführt werden. Ein Kernel ist ein Betriebssystemkern, indem Prozesse festgelegt sind, die die unterste Softwareschicht des Systems bilden und einen direkten Zugriff auf die Hardware haben.

Das Standard-Kernel ist IPython, ein Kommandozeilen-Interpreter, der das Arbeiten mit Python ermöglicht. [ionos]

Projekt Jupyter

In diesem Projekt wurden die drei Produkte Jupyter Notebook, JupyterHub und JupyterLab entwickelt, die sich hauptsächlich durch ihren Server voneinander unterscheiden.

Wie schon erwähnt wurde, profitiert Jupyter Notebook von einem Client-Server-Modell, während es sich bei JupyterHub um einen Multi-User-Server handelt. Ein Multi-User-Server dient dazu, dass mehrere Jupyter-Notebook-Instanzen miteinander verknüpft werden können. Von daher wird den Benutzern die Nutzung einer gemeinsamen Notebook-Umgebung ermöglicht.

Noch mehr Optionen zur Interaktion wird von JupyterLab, der offizielle Nachfolger von Jupyter Notebook, angeboten, indem es bei der Benutzeroberfläche zusätzliche Verknüpfungen zu Google Drive und weiteren Cloud-Diensten gibt.

🚩 Warum der Name Jupyter?

Der Name „Jupyter“ bezieht sich auf die drei wesentlichen Programmiersprachen Julia, Python und R.

Schema und Anwendungen in einem Beispielprogramm

Das Schema von Jupyter Notebook lässt sich hauptsächlich unter Input- und Outputzeilen beschreiben. Die Inputzeilen beinhalten die Eingabe bzw. den Code. Die Eingabe kann außerdem Kommentare, mathematische Formeln oder auch Gleichungen beinhalten. (siehe Fig. 2)

- ✓ Durch Kommentare wird der Code sowohl übersichtlicher als auch verständiger, für die Benutzer, die den Code möglicherweise bearbeiten möchten.

Das Ergebnis lässt sich in unterschiedlichen Formen nach dem Übersetzen des Codes in Outputzeilen zeigen. Auch als ein Diagramm kann der Benutzer sein gewünschtes Ergebnis graphisch gezeigt bekommen.

Der Typ der Inputzeilen ist eine Liste (list) von Strings, und im Gegensatz dazu sind die Outputzeilen vom Typ „dictionary“. Das liegt daran, dass manche Inputzeilen kein Ergebnis liefern, sondern diese haben eine Speicherfunktion.

```
In [1]: items = ["a", "b", "c"]
```

```
In [2]: items[1]
```

```
Out[2]: 'b'
```

```
In [3]: print(items)
        ['a', 'b', 'c']
```

```
In [4]: items[1:]
```

```
Out[4]: ['b', 'c']
```

Fig. 2: Screenshot der Input- und Outputzeilen - [scry]

Man sieht in der zweiten Figur bei der ersten Inputzeile, dass die drei Symbole „a,b,c“ in einem String gespeichert werden, die man in der dritten Inputzeile aufrufen und zeigen lässt.

Außer die Input- und Outputzeilen stellt Jupyter Notebook mehrere andere Angebote zur Verfügung.

Der Benutzer hat die Möglichkeit, ein vorgefertigtes Projekt unter „File“ hochzuladen, oder allgemein ein neues eigenes Projekt bauen. Unter die drei Menüs „Edit“, „Insert“ und „Cell“ kann man die Zeilen beliebig bearbeiten, die man sowohl alleine als auch in Abhängigkeit von anderen Zeilen übersetzen lassen kann. (siehe Fig. 3)

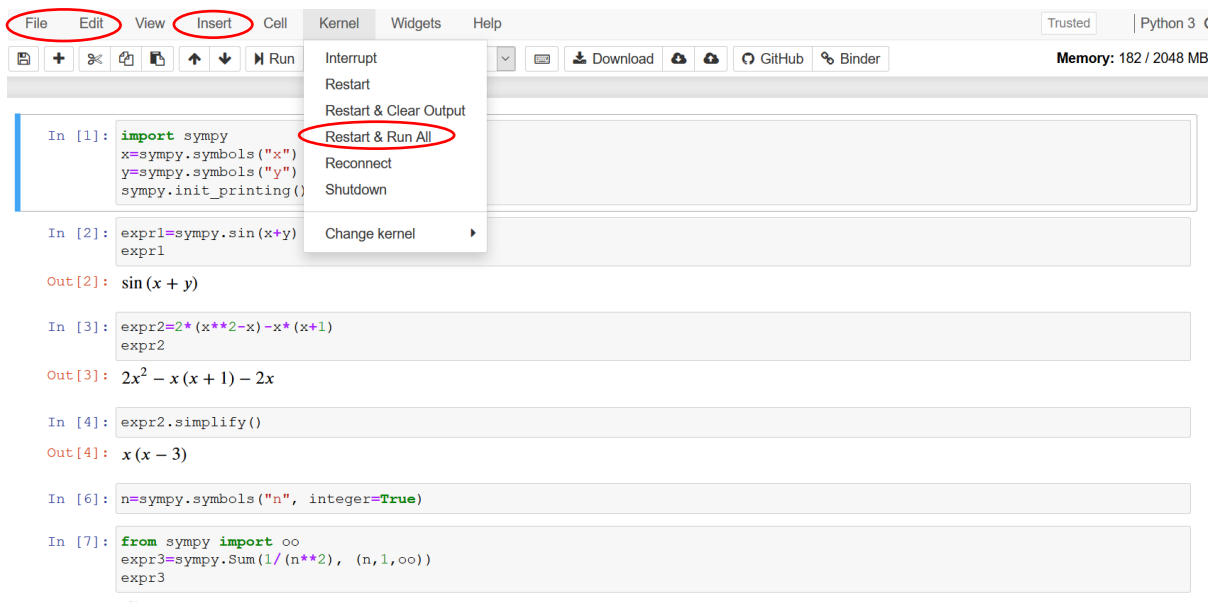


Fig. 3: Screenshot der Funktionen der Oberfläche

Hier wurde z.B. gezeigt, dass der Benutzer unter „Kernel“ das Programm neustarten und alle Zeilen gleichzeitig übersetzen kann. (siehe Fig. 3)

Aufgrund einer sehr vorteilhaften Funktion von Jupyter Notebook kann man direkt bei der Browseroberfläche den Code in anderen Formaten wie z.B. PDF oder Latex-Vorlagen herunterladen. (siehe Fig. 4)

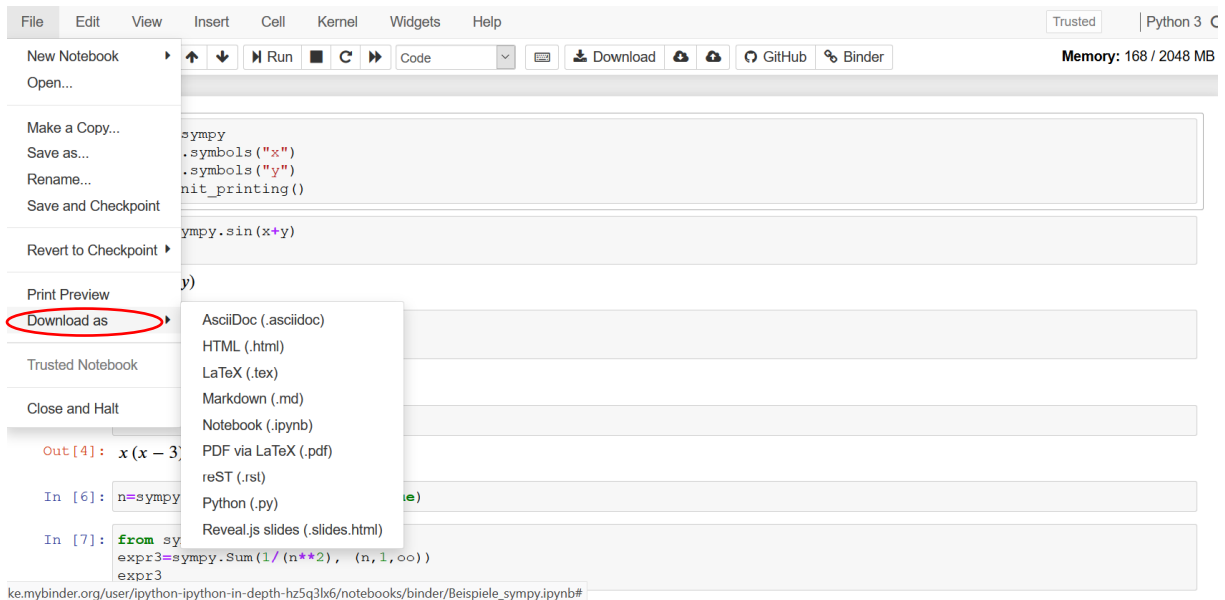


Fig. 4: Screenshot der Funktion „Download as“

Empfehlenswert für Anfänger ist es, dass die vorm Codieren eine Tour durch die Umgebung von Jupyter Notebook machen. Dies führt dazu, dass man zusammengefasst einen guten Überblick über den unterschiedlichen Möglichkeiten von der Oberfläche gewinnt. (siehe Fig. 5)

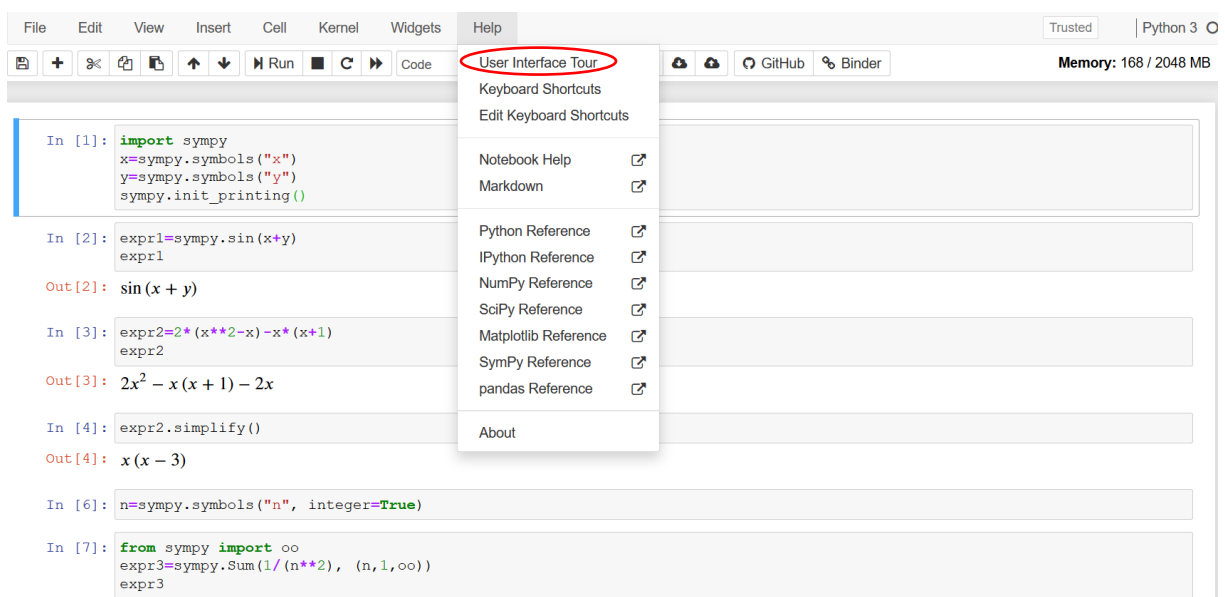
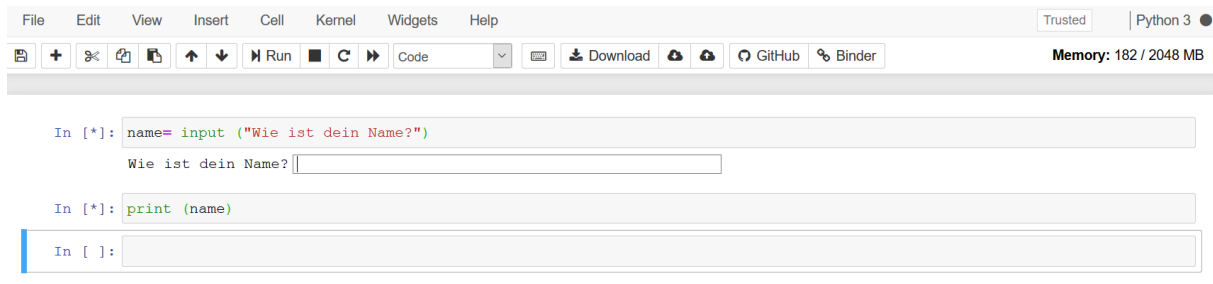


Fig. 5: Screenshot von „User Interface Tour“

Die Input- bzw. Outputzeilen werden in einer bestimmten Reihenfolge ausgeführt, sodass es keine Ausgabe geben wird, wenn deren Inputzeile vorher nicht übersetzt worden ist.



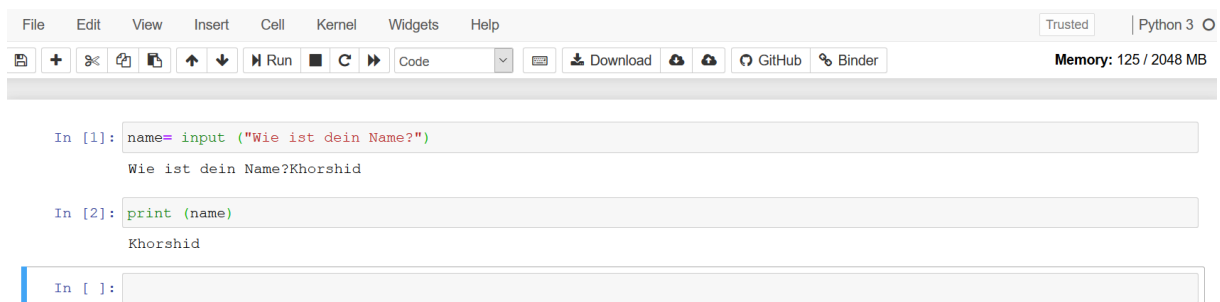
```
In [*]: name= input ("Wie ist dein Name?")
Wie ist dein Name? 

In [*]: print (name)

In [ ]:
```

Fig. 6: Screenshot eines Beispielprogramms (Schritt 1)

Alle Zeilen werden zwar komplett ausgeführt und man sieht die Nachricht „Wie ist dein Name?“, wo man den Namen eingeben sollte, aber es wird keine Ausgabe geben. Solange man keinen Namen reinput, wird man auch keinen Namen ausgedruckt bekommen. (siehe Fig. 6)



```
In [1]: name= input ("Wie ist dein Name?")
Wie ist dein Name?Khorshid

In [2]: print (name)
Khorshid

In [ ]:
```

Fig. 7: Screenshot eines Beispielprogramms (Schritt 2)

Nach der Übersetzung von der ersten Zeile und der Eingabe des Namens wird die zweite Zeile (In 2) erfolgreich ausgeführt. (siehe Fig. 7)

Funktion und Funktionalität

In diesem Teil geht es hauptsächlich darum, dass die zwei Begriffe „Funktion“ und „Funktionalität“ näher angeschaut und angedeutet werden.

Ein Programm besteht aus mehreren Funktionen, die nacheinander aber in Abhängigkeit voneinander laufen und ein einheitliches Ergebnis liefern. Das realisierte Ergebnis am Ende ist dann die Funktionalität des Programms.

Die grundlegende Funktionalität von Jupyter Notebook ist, dass der Benutzer einen fremden Code erstmal verstehen, verwenden und weiterentwickeln kann. Dazu kommen noch mehrere nützliche Eigenschaften, die hier kurz erwähnt werden [nwt]:

- ✓ Mit Hilfe von diesem Programm hat der Benutzer einen sehr einfachen und schnellen Zugriff auf einem fremden Code.
- ✓ Das Output nach dem Übersetzen des Codes ist vollkommen sichtbar.
- ✓ Der Programmierer kann für die bessere Übersichtlichkeit beliebig viele Kommentare hinzufügen, die sich im Quelltext zeigen lassen.
- ✓ Der Code kann direkt bei der Browseroberfläche in ein anderes Dateiformat konvertiert werden.
- ✓ Jupyter Notebook ermöglicht eine Datenvisualisierung, indem die Daten grafisch dargestellt werden können. Das Ergebnis bzw. das Output kann sogar in Form von Diagrammen graphisch dargestellt werden.
- ✓ Außerdem können mathematische Ausdrücke als Code eingegeben und definiert werden. Sogar komplexe mathematische Berechnungen lassen sich mit Hilfe von Jupyter Notebook lösen.

- ✚ Damit der Programmierer sein gewünschtes Ergebnis aus dem kompilierten Code bekommen kann, muss der Code einige Voraussetzungen erfüllen [nwt]:

Der Code muss richtig strukturiert und gut lesbar sein. Um den Code verständiger zu machen, ist es auch sinnvoller, dass der Code durch Kommentare dokumentiert wird.

Um der Zugriff auf dem Code zu vereinfachen, soll der Code über GitHub oder ähnliche Plattform versionisiert werden. Sodass das ganze Programm als Package verfügbar sein kann.

Die letzte Voraussetzung für den Code ist, dass der Code plattformunabhängig sein soll. Plattformunabhängigkeit ist eine Eigenschaft von einem Programm, das mit unterschiedlichen Hard- und Software-Konstellationen kompatibel ist, und dadurch unterschiedliche Betriebssysteme wie Windows, Unix oder Mac OS oder verschiedene Rechnerarchitekturen sowie Macintosh oder PC unterstützt. Ein plattformabhängiger Code ist außerdem überall lauffähig. [wiki3]

Vorteile

Einige wichtige Vorteile von Jupyter Notebook sind hier noch zu erwähnen [ionos]:

- ✓ Durch das Erstellen von Diagrammen, das auch mit Hilfe von Jupyter Notebook möglich ist, können die Daten in einer sehr einfachen Art visualisiert werden.
- ✓ Das Ausführen mathematischer Ausdrücke lässt sich hiermit von Jupyter Notebook in Echtzeit getestet werden.
- ✓ Diese Plattform unterstützt mehr als 50 verschiedene Programmiersprachen sowie Python, Java, PHP etc.
- ✓ Die Nutzbarkeit von diesem Programm ist kostenlos und läuft über die Browseroberfläche.
- ✓ Das JSON-Dokument, das dadurch erstellt wird, kann sehr einfach von der Oberfläche importiert werden.

Zusammenfassung

Abschließend kann man als Zusammenfassung folgende Punkte über Jupyter Notebook nennen:

- ✓ Jupyter Notebook ist eine Client Server Anwendung der Non Profit Organisation Project Jupyter.
- ✓ Jupyter Notebook ermöglicht die Kreation und das Teilen von Webdokumenten im JSON Format, die einem versionierten Schema und einer geordneten Liste von Input/Output Zellen folgen.
- ✓ Die erstellten Jupyter Dokumente lassen sich u. a. als HTML-, PDF-, Markdown- oder Python-Dokumente exportieren oder alternativ per E-Mail, Dropbox oder GitHub mit anderen Nutzern teilen.
- ✓ Sowohl mathematische Darstellungen als auch Visualisieren von Daten in Form von Diagrammen lassen sich mit Jupyter Notebook lösen.

Literaturverzeichnis

- [wiki1]: https://de.wikipedia.org/wiki/Project_Jupyter
- [ifros]: <https://www.ifross.org/was-open-source-software>
- [ionos]: <https://www.ionos.de/digitalguide/websites/web-entwicklung/jupyter-notebook/>
- [wiki2]: <https://de.wikipedia.org/wiki/Webanwendung>
- [nwt]: https://nwt-bw.de/wp-content/uploads/2020/01/PythonLehre_SMack_150220_final.pdf
- [code]: <https://www.codecademy.com/articles/how-to-use-jupyter-notebooks>
- [scry]: https://www.scrygroup.com/tutorial/2019-03-08/jupyter_cell_input_output/
- [wiki3]: <https://de.wikipedia.org/wiki/Plattformunabh%C3%A4ngigkeit>