J113037 CAS Datenjournalismus Basiskurstage

|  |  |  |
| --- | --- | --- |
| Datum |  | 06. September 2021 bis 17. September 2021 |
| Kurszeit |  | 09.15 Uhr bis 16.45 Uhr |
| Kursort |  | MAZ - Die Schweizer Journalistenschule, Murbacherstrasse 3, Luzern  Empfang und Zugang zu den Kursräumen im 1. Stock. |

|  |  |  |
| --- | --- | --- |
| Dozierende |  | Alexandra Stark, Studienleiterin, MAZ - Die Schweizer Journalistenschule  Thomas Ebermann, Data Scientist, Liip AG  Simon Schmid, Ressortleiter Wirtschaft, Wissenschaft, Digitales, Republik Project R Genossenschaft  Barnaby Skinner, Ressortleiter Visuals & Data , NZZ Mediengruppe |

|  |  |  |
| --- | --- | --- |
| Technical -Assistants |  | Mark Walther  Lea Senn, Fixxpunkt AG Rechnungseingang |
| Vorbereiten |  | 1. Schärft eure Ideen für drei konkrete datenjournalistische Daten-Projekte.   Und wie bereits im Willkommensmail erwähnt:   1. Arbeitet zwei Tutorials durch:  * [Python 3 Tutorial](https://www.sololearn.com/Course/Python/). Arbeitet die Module 1 und 2 durch. Es geht darum, ein Gefühl dafür zu bekommen, was Python überhaupt ist. Es geht hingegen nicht darum, am Ende der Module 1 und 2 bereits Python schreiben zu können. Das lernen wir im Kurs. * [Introduction to Statistics](https://github.com/MAZ-CAS-DDJ/kurs_19_20/blob/master/00%20weitere%C2%A0Dokumente/reading/Online_Statistics_Education.pdf) (schweres PDF, 100MB). Arbeitet Kapitel 1 durch, also genauer die Introduction bis Seite 64. Dieses Kapitel dient als Refresh für alle, die sich schon lange nicht mehr mit den Kernkonzepten der Statistik auseinandergesetzt haben; oder für jene, die das vielleicht nie getan haben.  1. Installiert Slack und die neuste Version von Google Chrome. |
| Mitbringen |  | 1. Laptop (wenn immer möglich Mac oder Linux) mit kompletten Admin-Rechten (Geschäftslaptop somit wenig geeignet). Idealerweise handelt es sich auch um ein Gerät, mit dem ihr bereits zu arbeiten gewohnt seid. 2. Allenfalls zusätzliche Zeit: Es kann hilfreich sein, die Abende für allfälliges Weiterarbeiten freizuhalten. |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| Lernziele |  | Die Teilnehmerinnen und Teilnehmer  \_wissen, welche Daten und welche Datenbanken für journalistisches Arbeiten interessant sind, wie eine datenjournalistische Story geplant, laufend überprüft und dokumentiert wird und können dies auf eigene Projekte anwenden;  \_können mit Python (inkl. diversen Zusatz-Modulen und -Libraries) eigene Programme schreiben, die es ihnen erlauben, Daten zu beschaffen, zu reinigen, zu strukturieren und zu analysieren;  \_ wissen, wie man mit Github arbeitet, Fehler in Programmiercodes findet (Debug-Strategien) und sich in der Community der Programmierer Hilfe holen kann (Stackoverflow);  \_ kennen die Grundbegriffe von HTML;  \_ können Daten von Websites scrapen (Beautifulsoup und Selenium);  \_wissen, wie sie einen eigenen Server bedienen, um Scripts automatisiert auszuführen;  \_ können mit dem Python-Visualisierungs-Modul Pandas Story-Thesen finden, überprüfen und justieren.  \_ haben anhand von praktischen Beispielen statistisches Know-How vertieft. |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Montag, 06.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Alexandra Stark Barnaby Skinner |  | **Themen finden und datenjournalistische Recherchen planen**  \_ Einführung: Was ist Datenjournalismus?  \_ Best- und Worst-Practice-Beispiele: Das storyorientierte Denken ist zentral – von Beginn weg.  \_ Wie findet man datenjournalistisch interessante Themen?  \_ Recherchemethodik im Datenjournalismus  \_ Arbeiten mit Spreadsheets  \_ Grundschritte des datenjournalistischen Arbeitens: Daten beschaffen, reinigen, strukturieren, analysieren; iterierender Loop: Programmieren, Analysieren, Visualisieren, These justieren  \_ Ressourcenplanung: Aufwand/Ertrag im Auge behalten  \_ Kennenlernen von Google Colab und Markdown  \_ Kennenlernen von Github Desktop | |
|  |  |  | |
| Dienstag, 07.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Thomas Ebermann Mark Walther |  | **Python: Grundbegriffe**  Die Teilnehmenden  \_ erhalten eine Einführung in Variables, data types, lists und dictionaries;  \_ lernen Befehle wie for Loop; while Loop; if/else/elif; functions; list comprehensions;  \_ erkennen, welche Programme und Befehle für welche Formen von Daten hilfreich sind (strukturierter Datensatz direkt auf Datenträger; unstrukturierter Datensatz auf Datenträger; offen im Internet abfragbarer Datensatz; nur durch Interaktion im Internet abfragbarer Datensatz)  \_ üben den datenjournalistischen Arbeitsablauf (Datenabfrage; Datenstrukturierung, Datenanalyse und Datenvisualisierung);  \_ lernen erste Strategien, wie sie Programmierfehler finden (len; print etc.) und wie sie Stackoverflow für die Problemlösung einsetzen können. | |
|  |  |  | |
| Mittwoch, 08.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Thomas Ebermann Mark Walther |  | **Python: Grundbegriffe**  Fortsetzung | |
|  |  |  | |
| Donnerstag, 09.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Thomas Ebermann Lea Senn |  | **Python: Grundbegriffe**  Fortsetzung | |
|  |  |  | |
| Freitag, 10.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Barnaby Skinner  Lea Senn |  | **Von der Cloud zur lokalen Installation**  Die Teilnehmenden  \_ durchlaufen das Anaconda-Setup  \_ lernen Jupyter Notebook kennen; dabei werden Unterschiede und die Ähnlichkeiten zu Google Colab herausgearbeitet  \_ legen sich ganz bewusst einen Bedienungsprozess des eigenen Computers zuecht  \_ lernen was Python Libraries und Module sind  \_ lernen, wie mit dem Module os gearbeitet werden kann, um grosse Datenmengen ein- oder auszulesen  \_ erhalten einen Ausblick auf Requests und der Arbeit mit APIs  \_ Vertiefen ihre Kenntnisse von Github und legen sich auch hier einen für sie sinnvollen Arbeitsprozess zurecht | |
|  |  |  | |
| Montag, 13.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Barnaby Skinner Mark Walther |  | **Requests, APIs, Scraping mit Python (Beautifulsoup)**  Die Teilnehmenden  \_ erhalten einen Überblick über die Scraping Module Beautiful Soup 4, lxml, Scrapy;  \_ zeichnen mit HTML eine erste Scraping-Übung aus und lösen ein selbst konstruiertes HTML-Beispiel;  \_ erkennen typische Bauweisen von Websites, die fürs Scraping genutzt werden können;  \_ üben ausgewählte Funktionalitäten der verschiedenen Libraries;  \_ erkennen, für welche Formen von Daten Beautifulsoup hilfreich ist;  \_ fügen die neuen Fertigkeiten in den datenjournalistischen Arbeitsablauf ein und überlegen, was die Programme ihren Projekten bringen;  \_ lernen weitere Debugging-Strategien kennen und setzen Stackoverflow für die Problemlösung ein. | |
|  |  |  | |
| Dienstag, 14.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Barnaby Skinner |  | **Requests, APIs, Scraping mit Python (Beautifulsoup)**  Fortsetzung | |
|  |  |  | |
| Mittwoch, 15.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Simon Schmid Mark Walther |  | **Daten-Analyse und Visualisierung mit Python (Pandas)**  In diesem Kursteil sehen wir uns Pandas an, eine Statistik-Library, die für Börsenmakler entwickelt wurde. Mittlerweile aber von Wissenschaftlern aus der ganzen Welt für ihre unterschiedlichsten Forschungen eingesetzt wird.  Die Teilnehmenden...  - erhalten eine Einführung in Pandas;  - arbeiten die Grundlagen der Statistik auf;  - analysieren eigene gescrapte Datensätze;  - kombinieren Datensätze mit matplot lib, um sie zu visualisieren;  - automatisieren den eigenen Grafik-Stil;  - fügen Pandas in den datenjournalistischen Arbeitsablauf ein und überlegen, was das Programm ihren Projekten bringt. | |
|  |  |  | |
| Donnerstag, 16.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Simon Schmid Lea Senn |  | **Daten-Analyse und Visualisierung mit Python (Pandas)**  Fortsetzung | |
|  |  |  | |
| Freitag, 17.09.2021 |  | 09.15 bis 16.45 Uhr | Hauptraum 11 |
| Alexandra Stark Thomas Ebermann Lea Senn Barnaby Skinner |  | **Abschluss mit Selenium und Ausblick auf die Projekte**  \_ Selenium: HTML-Daten scrapen in Interaktion mit der Website;  \_ Übungen zu allen in den zehn Tagen gelernten Skills von Python-Functions über Requests, Beautifulsoup, Pandas bis zu Selenium.  \_ Planung der eigenen Projekte; Präsentation;  \_ Ausblick auf die neun Vertiefungstage, die Abschlussarbeit und den Abschlusstag.  \_ Feedbackrunde | |
|  |  |  | |