

Gültig für 2024.FS

**Modulbezeichnung: Python-Grundlagen & Anwendung in Data Science**

Modulkürzel	w.BA.XX.2PythGr.XX	
Beschreibung des Moduls	<p>Python hat als objektorientierte Allzweck-Programmiersprache in den letzten Jahren hinsichtlich der Beliebtheit und Verbreitung andere Programmiersprachen aufgeholt und wird dadurch zunehmend wichtiger. Dieses Modul bietet einerseits einen Überblick über die Anwendungsbereiche für Python und vertieft andererseits dessen Anwendung im Bereich des Data Science. Die Studierenden erhalten Code-Vorlagen für gängige Anwendungen, Methoden, Skripts und Modelle. Weiter wenden die Studierenden ihre Python-Kenntnisse auf eigenen Projekten an. Sie lernen Methoden und Werkzeuge, mit deren Hilfe sie ihr Wissen selbstständig in zugehörigen Bereichen weiterentwickeln können.</p>	
Studiengang und Vertiefungsrichtung	<p>§ Wirtschaftsinformatik § Wirtschaftsinformatik - Vertiefung in Business Information Systems</p>	
Rechtliche Grundlagen	<p>Studiordnung BSc vom 29.01.2009 für die Bachelorstudiengänge Betriebsökonomie, International Management, Wirtschaftsinformatik, Wirtschaftsrecht und Angewandtes Recht, erstmals beschlossen am 12.05.2009</p>	
Modulkategorie	<b>Modultyp:</b> Wahlpflichtmodul	<b>Studienabschnitt:</b> Hauptstudium
ECTS	3	
Verantwortliche OE	W Institut für Wirtschaftsinformatik	
Modulverantwortung	Maria Pelli (pell)	
Modulverantwortung Stellvertretung	Pasquale Cirillo (ciri)	
Spezifische Vorkenntnisse	<p>Grundkenntnisse für Machine Learning - Modelle (bspw. Data Analytics oder Machine Learning I) werden vorausgesetzt. Erste Programmiererfahrungen sind von Vorteil.</p>	
Beitrag des Moduls für Learning Goals des Studiengangs (durch das Modul betroffene)	<p>Fachkompetenz § Theorie- &amp; Praxisrelevante Fachinhalte wissen &amp; verstehen § Theorie- &amp; Praxisrelevante Fachinhalte anwenden, analysieren und verknüpfen § Theorie- &amp; Praxisrelevante Fachinhalte evaluieren</p> <p>Methodenkompetenz § Problemlösung &amp; Kritisches Denken § Wissenschaftliche Methoden § Arbeitsmethoden, -techniken &amp; -verfahren § Nutzung von Informationen § Kreativität &amp; Innovation</p> <p>Sozialkompetenz § Schriftliche Kommunikation § Mündliche Kommunikation § Kooperation im Team &amp; Umgang mit Konflikten § Interkulturalität &amp; Perspektivenübernahme</p> <p>Selbstkompetenz § Selbstmanagement &amp; Selbstreflexion § Ethische &amp; Soziale Verantwortung § Lernen &amp; Veränderung</p>	
Lernziele des Moduls	<p>Die Studierenden...</p> <ul style="list-style-type: none"> <li>§ kennen und verstehen die Grundlagen und Einsatzgebiete von Python</li> <li>§ verstehen den grundlegenden Aufbau eines Python-Programmes</li> <li>§ kennen Datenformate und können Datenströme selbst implementieren</li> <li>§ sind in der Lage geeignete Bibliotheken zu wählen und anzuwenden (u.a. Pandas, Matplotlib und Seaborn)</li> <li>§ kennen den grundlegenden Prozess der Datenaufbereitung, -bereinigung, -visualisierung und modellierung (Regressions- und Klassifikationsmodelle) und können diesen selbstständig auf Fragestellungen anwenden</li> <li>§ sind in der Lage, Lösungen für Fragestellungen im Bereich des Data science zu identifizieren und zu evaluieren</li> <li>§ können einfache Prototypen mit Python selbst entwickeln</li> <li>§ werden befähigt, ihr Wissen in den zugehörigen Bereichen selbstständig weiterzuentwickeln</li> </ul>	

Inhalte des Moduls	<ul style="list-style-type: none"> <li>§ Einführung in Python und dessen Anwendung in Data Science</li> <li>§ Grundstruktur eines Python-Programmes und dessen Elemente</li> <li>§ Kontrollstrukturen, Funktionen, Methoden und Attribute</li> <li>§ Objektorientierte Programmierung vs. Scripting in Python</li> <li>§ Datentypen- und formate, Data Frames, Series</li> <li>§ Methoden der Code-Optimierung (u.a. List Comprehensions, Alternativen für Schleifen, Funktionen und "built-ins")</li> <li>§ Explorative Datenanalyse (EDA) und Visualisierung der Daten (u.a. Pandas, Matplotlib und Seaborn)</li> <li>§ In- und Export von Daten und Datenströme (u.a. Pandas)</li> <li>§ Datenspeicherung für strukturierte, semi-strukturierte und unstrukturierte Daten (u.a. SQLite, MySQL, MariaDB, optional PyMongo)</li> <li>§ Filterung, Sortierung, Verbindung und Aggregation (u.a. Pandas)</li> <li>§ Duplikate, Dubletten, Fehlende Daten, Kodierung und Imputation (u.a. Pandas, NumPy, Scikit-Learn)</li> <li>§ Deskriptive, univariate und multivariate Analyse, Ausreisser, Validierung der Daten (u.a. Pandas, NumPy, Scikit-Learn)</li> <li>§ Feature Engineering von kategorialen und quantitativen Variablen (u.a. Pandas, Scikit-Learn)</li> <li>§ Standardisierung von Verteilungen, Normalisierung auf Wertebereiche, Hyperparameter und Quantizierung (u.a. Pandas, NumPy, Scikit-Learn, Tensorflow)</li> <li>§ Supervised Learning (u.a. Scikit-Learn, Tensorflow)</li> <li>§ Explainable AI - Methoden (u.a. SHAP)</li> <li>§ Prototyping und Deployment-Ansätze (u.a. Widgets, Dash, Flask)</li> </ul>																												
Verknüpfung zu anderen Modulen	-																												
Unterrichtsmethoden	<ul style="list-style-type: none"> <li>§ Lehrvortrag</li> <li>§ Übungen</li> <li>§ Projektarbeit</li> <li>§ Literaturstudium</li> </ul>																												
Digitale Lernressourcen	<ul style="list-style-type: none"> <li>§ Übungs- und Anwendungsaufgaben (inkl. Lösungen)</li> <li>§ Fallstudien (inkl. Lösungen)</li> <li>§ Beispielcodes in Python</li> </ul>																												
Unterrichtsgliederung	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th><th><b>Kontaktstudium</b></th><th><b>Begleitetes Selbststudium</b></th><th><b>Autonomes Selbststudium</b></th></tr> </thead> <tbody> <tr> <td>Grossklasse</td><td>-</td><td>-</td><td></td></tr> <tr> <td>Kleinklasse</td><td>28 h</td><td>-</td><td></td></tr> <tr> <td>Gruppenunterricht</td><td>-</td><td>-</td><td></td></tr> <tr> <td>Praktikum</td><td>-</td><td>-</td><td></td></tr> <tr> <td>Seminar</td><td>-</td><td>-</td><td></td></tr> <tr> <td><b>Total</b></td><td><b>28 h</b></td><td><b>0 h</b></td><td><b>62 h</b></td></tr> </tbody> </table>		<b>Kontaktstudium</b>	<b>Begleitetes Selbststudium</b>	<b>Autonomes Selbststudium</b>	Grossklasse	-	-		Kleinklasse	28 h	-		Gruppenunterricht	-	-		Praktikum	-	-		Seminar	-	-		<b>Total</b>	<b>28 h</b>	<b>0 h</b>	<b>62 h</b>
	<b>Kontaktstudium</b>	<b>Begleitetes Selbststudium</b>	<b>Autonomes Selbststudium</b>																										
Grossklasse	-	-																											
Kleinklasse	28 h	-																											
Gruppenunterricht	-	-																											
Praktikum	-	-																											
Seminar	-	-																											
<b>Total</b>	<b>28 h</b>	<b>0 h</b>	<b>62 h</b>																										
Leistungsnachweise																													
<b>Modulendprüfung</b>	<b>Form</b>	<b>Dauer (Min.)</b>	<b>Gewichtung</b>																										
-	-	-	-																										
<b>Hilfsmittel</b>	-																												
<b>Andere</b>	<b>Bewertung</b>	<b>Dauer (Min.)</b>	<b>Gewichtung</b>																										
Projektarbeit	Note	-	70,00 %																										
Referat/mündliche Präsentation	Note	20	30,00 %																										
Präsenzverpflichtung im Kontaktstudium	Zwingende Präsenzzeit: Keine																												
Unterrichts- und Prüfungssprache	Deutsch																												
Pflichtliteratur	McKinney, W. (2022). Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. 3. Auflage. O'Reilly Media. ISBN 978-1098104030. <a href="https://www.oreilly.com/library/view/python-for-data/9781098104023/">https://www.oreilly.com/library/view/python-for-data/9781098104023/</a> .																												
Ergänzende Literatur	James, G., Witten, D., Hastie, T., Tibshirani, R. & Taylor, J. (2023). An Introduction to Statistical Learning with Applications in Python. Springer Cham. ISBN 978-3-031-38747-0. <a href="https://drive.google.com/file/d/1ajFkHO6zjrdGNqhqW1jKBZdiNGh_8YQ1/edit">https://drive.google.com/file/d/1ajFkHO6zjrdGNqhqW1jKBZdiNGh_8YQ1/edit</a> .																												
Bemerkungen	Es werden ein eigenes Notebook sowie ein eigenes Google Colab - Konto vorausgesetzt.																												