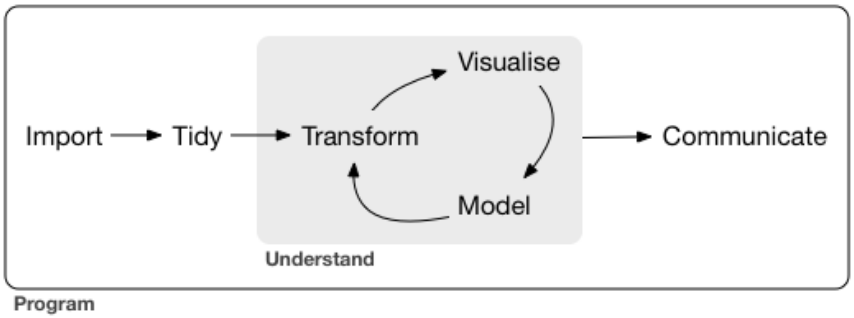
Modulname	**Data Analytics**			
Stand	September 2020			
Modulcode	DAAN			
Anzahl ECTS-Punkte	4			
Gesamtarbeitsaufwand / Workload in Stunden	1 ECTS entspricht 30h Aufwand für die Studierenden jede Lektion (1h Kontaktstudium/Woche) ergibt 14h/Semester			
	Kontaktstudium **Anteil Theorie** **und Übung** **Praktikum (P)** (Kleingruppen)	davon V: 2 Lekt. Ü: 2 Lekt. P: Lekt.	Total: 28 Stunden 28 Stunden Stunden	
	Begleitetes Selbststudium	„Projekt-Arbeiten“	Stunden	
	Individuelles Selbststudium	Prüfungs-Vorb.	64 Stunden	
	Total		**120 Stunden**	
Regel-Semester	**Vollzeit:** 5. Semester	**Teilzeit:** 7. Semester		
Unterrichtssprache	Deutsch			
Modulniveau (Erklärung am Ende)	☐ B	☐ I	☒ A	☐ S
Modultyp (Erklärung am Ende)	☐ C Pflicht	☐ R Stand.	☒ M Wahl	
Modulverantwortliche(r)	Adrian Stämpfli			
Dozierende	Adrian Stämpfli			
Lehr-/Lernmethoden primäres Konzept	Lehrgespräch, Referat, Einzel-, Partner-, Gruppenarbeit, Selbststudium			

Leitidee der Umsetzung	Das Modul "Data Analytics" vermittelt den Studierenden eine solide Basis in allen Teilschritten typischer Datenanalysen. Es orientiert sich dabei an folgendem Ablaufdiagramm:  Quelle: Wickham and Grolemund 2017, https://r4ds.had.co.nz/introduction.html#what-you-will-learn Die Studierenden lernen jeden Teilschritt exemplarisch kennen und üben die Teilschritte mit Übungsaufgaben. Als Tool kommt jeweils R/RStudio zum Einsatz.
Praxisbezug Sicherstellung Praxisbezug	
Umsetzung der WING-Anf Kommunikation Teamarbeit Systemdenken Industrielle Prozesse Mechatronik	
Angestrebte Lernergebnisse (Abschlusskompetenzen)	**Fachkompetenzen:** Die Teilnehmenden können: • Datensätze nach R importieren. • Datensätze in ein konsistentes Datenformat überführen. • Daten mithilfe geeigneter R-Funktionen in saubere, konsistente und praktische Formate überführen. • Datensätze explorativ erschliessen mittels Datenvisualisierung und Hypothesengenerierung/-bestätigung. • Validierungsverfahren (z.B. Residuen-Analyse und Kontrolle auf Overfitting mittel Train- und Test-Set. • Einfache statistische Modelle (Lineare Regression, Clustering) in der Praxis anwenden. • überwachte (supervised) Modelle zur Hypothesengenerierung einsetzen. • Datenanalysen mit Tools (z.B. RMarkdown, Shiny) auf einfachste Weise kundengerecht aufbereiten und zugänglich machen.

	Methodenkompetenzen: Die Teilnehmenden können: • zentrale Ergebnisse aus Datenanalysen in einer konkreten Fallstudie kommunizieren. • mit fehlenden Datensätzen und Ausreissern umgehen. • Datensätze transformieren (Variablen zusammenführen, neue Variablen berechnen, filtern).
	Selbstkompetenzen: Die Teilnehmenden können: • Eine relativ komplexe Datenanalyse Aufgabe sinnvoll in Teilschritte zerlegen und zielgerichtet umsetzen.
	Sozialkompetenzen: Die Teilnehmenden können: • Datenanlysen von Kolleg_innen sinnvoll challengen und Feedback gezielt anbringen

Modul-/Lerninhalte	**Themen-/Lernblock:** • **Import/Data-Cleaning/Data-Preparation:** Die Studierenden lernen Datensätze nach R zu importieren und in ein konsistentes Datenformat zu überführen (Aus Files wie Excel und CSV, aus APIs sowie aus Datenbanken). Sie können mit fehlenden Datensätzen und Ausreissern umgehen. Sie können Daten mithilfe geeigneter R-Funktionen in saubere, konsistente und praktische (für die Weiterverarbeitung) Formate überführen. • **Transform/Visualise/Model:** Die Studierenden lernen den iterativen Prozess "from Data to Meaning" praktisch kennen. Die Teilschritte "transform → visualise → model" werden meist mehrfach durchlaufen. Im Einzelnen lernen die Studierenden: ○ Datensätze zu transformieren (Variablen zusammenführen, neue Variablen berechnen, Filtern etc.); Wir werden dabei sowohl "manuelle" Verfahren nutzen (gruppieren, summieren, normalisieren), als auch die PCA (Principal Component Analysis) exemplarisch einführen (wir werden dabei nicht mathematisch in die Tiefe gehen..). ○ Datensätze explorativ zu erschliessen; Hierzu gehört die Datenvisualisierung zum Zweck der "Erkundung" und die Hypothesengenerierung (und teilweise auch Bestätigung). Wir nutzen aber auch Validierungsverfahren wie z.B. die Residuen-Analyse und Kontrolle auf Overfitting mittels Train- und Test-Set. ○ Einfache statistische Modelle in der Praxis anwenden (Lineare Regression, Clustering). Sie lernen dabei, dass auch überwachte (supervised) Modelle zur Hypothesengenerierung eingesetzt werden können. • **Communicate:** Die Studierenden üben die Kommunikation zentraler Ergebnisse aus Datenanalysen in einer konkreten Fallstudie. Die Studierenden lernen dabei Tools (RMarkdown, Shiny) kennen, welche es Ihnen erlauben ihre Datenanalysen auf einfachste Weise kundengerecht aufzubereiten und zugänglich zu machen.
Vorkenntnisse (Eingangskompetenzen)	• Die Studierenden haben den einwöchigen Blockkurs "Programming for Data Science" absolviert. Dies garantiert die notwendigen Vorkenntnisse in Programmierung, dem Umgang mit Daten und im Speziellen mit der Umgebung R/RStudio. • Die Studierenden haben Vorkenntnisse in Statistik. • Optimalerweise haben die Studierenden Vorkenntnisse in Modellbildung (MOSIM).
Lehrmittel/-materialien **Methoden** **Vorlagen**	Pflichtliteratur: (Skript, Bücher, etc) - R for Data Science (Wickham & Grolemund 2017)

Konzepte	Weiterführende Literatur: *R for Data Science* (Wickham & Grolemund 2017); *ggplot2* (Wickham 2016); *R Programming for Data Science* (Peng 2016)	
Leistungsnachweise: **Prüfungsart und -dauer**	☐ schriftliche Prüfung ☒ Präsentationen, Dauer: 30 Minuten ☐ Korreferate ☐ Projekte ☐ Lernberichte ☒ schriftliche Arbeiten: Durchführung eines Data Science Projekts ☐ andere, nämlich:	
Leistungsnachweise: Weitere Angaben (z.B. Gewichtung der Prüfungsteile bei mehreren Leistungsnachweisen, erlaubte Hilfsmittel, Anforderungen	Die Studierenden führen ein konkretes Data Science Projekt durch, erstellen einen schriftlichen Schlussbericht (mit R/RStudio und markdown) und präsentieren die Ergebnisse in einer Präsentation.	
Zulassungsbedingungen zu den Prüfungen	Keine	
NICHT enthaltene Inhalte werden explizit in einem anderen Modul erwartet oder vorausgesetzt!		
Inhalte Industrieprojekt werden explizit im Industrieprojekt behandelt!	1 - Potential	
	2 - Idee	
	3 - Entwurf	
	4 - Prototyp	
	5 - Optimierung	
Geplante Bildungsausflüge Exkursionen, Firmenbesuche	Keine	
Notwendige Systeme Software, Hardware Ausrüstung Zimmer Praktika etc. (Investitions-Planung)	R/R Studio	
Besonderes		

Legende Modulniveau:
B – Basic level course: Modul bzw. Kurs zur Einführung in das Basiswissen eines Gebiets

I – Intermediate level course: Modul bzw. Kurs zur Vertiefung der Basiskenntnisse
A – Advanced level course: Modul bzw. Kurs zur Förderung und Verstärkung der Fachkompetenz¨
S – Specialised level course: Modul bzw. Kurs zum Aufbau von Kenntnisse und Erfahrungen in einem Spezialgebiet

Legende Modultyp:
C – Core course: Modul bzw. Kurs des Kerngebiets eines Studienprogramms (Pflichtmodul bzw. Pflichtkurs)
R – Related course: Unterstützungsmodul bzw. -kurs zum Kerngebiet (z.B. Vermittlung von Vor- oder Zusatzkenntnissen) (Wahlpflichtmodul bzw. -kurs)
M – Minor course: Wahlmodul bzw. -kurs