Studiengang: B.Sc. Maschinenbau und Produktion B.Sc. Maschinenbau und Produktion (dual)	
Modulbezeichnung / Titel **Module name / title (engl.)**	**Big Data** **Big Data**
Modulkennziffer	BDA
Modulkoordination/ Modulverantwortliche/r	Frau Prof. Dr. Sarah Hallerberg
Dauer des Moduls/ Semester/ Angebotsturnus	1 Semester/ 5. oder 6. Semester, im dualen Studiengang 6. oder 7. Semester/ jährlich
Leistungspunkte(LP)/ Semesterwochenstunden(SWS)	5 LP/ 4.00 SWS
Art des Moduls, Verwendbarkeit des Moduls	Wahlpflichtfach in der Studienrichtung Digital Engineering and Mobility Wahlpflichtfach in dem Studienschwerpunkt Service Engineering
Arbeitsaufwand (Workload)	Präsenzstudium 68 h und Selbststudium 82 h (17 Semesterwochen, 1 SWS = 60 min)
Teilnahmevoraussetzungen/ Vorkenntnisse	Fehlen Prüfungs- oder Studienleistungen des 1. und 2. Semesters, können keine Prüfungsleistungen ab dem 5. Semester abgelegt werden. Empfohlen: Mathematik, Angewandte Informatik
Lehrsprache	Regelhafte Lehrsprache: Deutsch Weitere mögliche Lehrsprache: Englisch Bei mehr als einer möglichen Lehrsprache im Modul wird die zu erbringende Lehrsprache von dem bzw. der verantwortlichen Lehrenden zu Beginn der Lehrveranstaltung bekannt gegeben.
Zu erwerbende Kompetenzen/ Lernergebnisse	Digitalisierung und Industrie 4.0 erfordern den sicheren Umgang mit Datenmengen aus Sensoren, Kundenanforderungen oder anderen Quellen. Zielkompetenzen dieses Moduls: Die Studierenden können große Datensätze speichern und analysieren. Sie kennen Hardware-Systeme, Dateisysteme und Datenformate für die Verarbeitung großer Datenmengen und können mit diesen arbeiten. Des Weiteren kennen sie wichtige Algorithmen, die die Analyse großer Datenmengen vereinfachen. Sie können dieses Algorithmen verstehen und anwenden. Des Weiteren können sie eigene Algorithmen so erstellen, daß die Verarbeitung großer Datenmengen beschleunigt wird. Hierzu erwerben die Studierenden innerhalb der Vorlesungen theoretisches Hintergrundwissen und wenden dies in den Laboren in (Gruppen-) projekten an.

Inhalte des Moduls	1. Einführung in Programmiersprachen und Software, Verteilung der Projektdatensätze 2. Relationelle Datenbanken, No-SQL-Datenbanken, Graphenorientierte-Datenbanken, Distributed-Ledger-Technologie, Anwendungstests an Teilen der Projektdatensätze 3. hochverfügbare verteilte Dateisysteme zum Aufbau eines Data-Warehouses, Anwendung an Teilen der Projektdatensätze 4. Werkzeuge zur Konfiguration und Überwachung eines Data Warehouses, Tests an Teilen der Projektdatensätze und am hadoop-Cluster der HAW 5. Map Reduce und weiterführende Ansätze, Idee, Funktionsweise und Anwendungstests an Teilen der Datensätze 6. „Lambda-Kalkül“, Funktionsweise und Anwendung an Teilen der Projektdatensätzen 7. Schnittstellen und Event-Processing, Tests an Teilen der Projektdatensätze 8. Machine-Learning und Big Data, Tests an Teilen der Projektdatensätze 9. Bearbeitung der Projektaufgaben an den vollständigen Datensätzen (auf dem hadoop-Cluster der HAW)
Voraussetzungen für die Vergabe von Leistungspunkten (Studien- und Prüfungsleistungen)	Regelhafte Prüfungsform für die Modulprüfung: Projekt (PL) Laborpraktikum: Laborabschluss (SL)
Lehr- und Lernformen/ Methoden/ Medienformen	Seminaristischer Unterricht (3 SWS), Projektarbeiten im Labor (1 SWS) Tafel, Beamer, Computerpool, Laptopwagen
Literatur	D. Fasel, F. Meier, Big Data Grundlagen, Systeme und Nutzungspotenziale, Springer, 2016 T. White, Hadoop, the Definitive Guide, O' Reighly, 2015 https://spark.apache.org/docs/latest/api/python/index.html http://hadoop.apache.org/docs/stable/