**Data Science in Finance**

*Vorläufige Struktur*

**1. Block: Python Basics I**

* Programmiersprachen – Einführung und Überblick
  + Was ist eine Programmiersprache und welche wesentlichen Programmierparadigmen gibt es?
  + Was zeichnet Python aus und warum ist es derzeit die beliebteste Programmiersprach in Forschung und Praxis?
  + In welchen Aufgabenbereichen der Finanzwirtschaft kommt Python zum Einsatz?
* **Impulsvortrag**: Kurze Vorstellung AFC Modelling
  + Kurze Vorstellung AFC-M – Content TBD
  + Anwendung von Python innerhalb von AFC-M
* Installation und Einrichtung notwendiger Software
  + Python
  + Visual Studio Code
  + Git
* Grundlagen – Rudimentäre Operationen in Python
  + Arithmetische Rechenoperationen in einer Shell
  + Deklaration von Variablen und erste Datentypen
  + Wenn-Kondition und logische Operatoren
  + Einführung in Loops in Python
  + Erste Programmierübung im Kurs (TBD)
* Hausaufgabe I
  + TBD
  + TBD
  + TBD
  + Bonus Aufgabe (Hoher Schwierigkeitsgrad, Bonuspunkte, TBD)

**2. Block: Python Basics II**

* Diskussion Hausaufgabe I – Musterlösung und typische Fehler
* Funktionen – Eine erste Einführung
  + Definition einer Funktion: Parameter und Resultat
  + Lambda Funktionen – Als Kurzschreibweise von Funktionen
  + Beispiele von rekursiven Funktionen
* Klassen und Objekte
  + Eine Klasse als Template, Initialisierung und Methoden
  + Objekte als Instanz einer Klasse
  + Vererbung
* Regular Expressions
  + Buchstaben, Zahlen und Satzzeichen
  + Regex Anker
  + Regex Quantifizierungen
  + ODER-Kondition
  + Regex Klassen
  + Pattern Matching
  + Pattern Substitute
  + Gruppierungen
  + Klammer Operatoren
  + Wortgrenzen
  + Lazy und Greedy Match
* Hausaufgabe II
  + Country Code Derivation
  + TBD
  + TBD
  + Bonus Aufgabe (Hoher Schwierigkeitsgrad, Bonuspunkte, TBD)

**3. Block: Wichtige Python Bibliotheken für Data Science**

* Diskussion Hausaufgabe II – Musterlösung und typische Fehler
* Libraries – Grundlagen
  + User Defined Libraries
  + Importieren von Libraries
  + Installation und Import von Standard Libraries
  + Virtual Environments and Versioning
* Numpy
  + Arithmetische Operationen
* Pandas
  + Erstellen von Dataframes
  + Auswahl von Reihen, Spalten und Zellen
  + Manipulation von Zellen
  + Standardrechenoperationen
* Matplotlib
  + Balken- und Kreisdiagramme
  + Darstellung von Zeitreihen
* Hausaufgabe III
  + TBD
  + TBD
  + TBD
  + Bonus Aufgabe: Parsen von SWIFT Nachrichten

**4 Block: Fortgeschrittene Anwendungen**

* Diskussion Hausaufgabe III – Musterlösung und typische Fehler
* Git – Eine Einführung in die Versionsverwaltung
* Natural Language Processing – Motivation und Einführung
  + Wörterbuchbasierte Sentimentanalyse
  + Ähnlichkeitsanalysen
  + Named Entity Recognition
  + Topic Modelling
* Tensorflow – Google Machine Learning Library
  + Logistische Regression
  + Anwendung eines neuronalen Netzes
  + Machine Learning mit tensorflow
* Ausblick
  + Cloud Anwendungen – kurze Vorstellung von der Google Cloud Platform
  + Weitere Vertiefungsmöglichkeiten – Webapplikationen
* Q&A
* Projektarbeit in Gruppenarbeit
  + **Thema 1**: Parsen von OSM Kartendaten
    - Automatisierte Erstellung von Referenzlisten
    - Verwendung u.a. in der Country Code Derivation
    - Verbesserung des Country Codes durch Einbindung weiterer Features
  + **Thema 2:** TBD
  + **Thema 3:** TBD
  + **Thema 4:** TBD

**Weitere Informationen:**

* Pro Block ca. 270 min (6 Semesterwochenstunden)
* Hausaufgaben: 2 Wochen Bearbeitungszeit; maximal 1-2 Personen pro Gruppe
* TBD: Praxisvorträge

**Prüfungsleistung:** Projektarbeit

* Zufällige Verteilung drei verschiedener Projektarbeiten
* Bereitstellung umfangreicher Datensätze (Bspw. Country Code Test Daten, Swift-Test Daten, Kartegrafische Datenbanken, E-Mails, SEC Filings etc.)
* Erstellung von Reports (max. 10 Seiten) auf Basis umfangreicher Analysen
  + Kurze Motivation und Interpretation der Analyse
  + Visualisierung der Daten
  + Abgabe des zugrundeliegenden Python Codes
* Die Benotung setzt einen Fokus auf
  + das analytische Vorgehen (Code)
  + die sinnvolle Auswahl an Analysen
  + die adäquate Anwendung ausgewählter Data Science Tools (Code)
  + eine klare Aufbereitung des zugrundeliegenden Codes
  + die Gesamtschau des finalen Reports
* Die Projektarbeit ist in Einzelarbeit anzufertigen
* Die Bearbeitungsdauer umfasst ca. 4 Wochen