

Prof. Dr. Pascal Kerschke
Lennart Schäpermeier, M. Sc.

TU Dresden, Lehrstuhl für Big Data Analytics in Transportation

Spezifische Aspekte von Data Analytics

Winter Term 2021/2022

Spezifische Aspekte von Data Analytics: Kursorganisation



Prof. Dr. Pascal Kerschke

pascal.kerschke@
tu-dresden.de



Lennart Schäpermeier, M. Sc.

lennart.schaepermeier@
tu-dresden.de

Allgemeiner Ablauf:

- Zusammenarbeit in kleinen Gruppen von 2-3 Studierenden
- Projekte bestehen aus der Anwendung eines von mehreren vorgeschlagenen Themen
- Benotung basierend auf: Zwischenpräsentation, Abschlusspräsentation und Seminarbericht

Voraussetzungen:

- dies ist ein *Fünftsemesterkurs*
- Sie sollten den Kurs *Data Analytics Grundlagen* bereits (erfolgreich) absolviert haben
- grundlegende Programmierkenntnisse in R oder Python

Allgemeiner Ablauf:

- Zusammenarbeit in kleinen Gruppen von 2-3 Studierenden
- Projekte bestehen aus der Anwendung eines von mehreren vorgeschlagenen Themen
- Benotung basierend auf: Zwischenpräsentation, Abschlusspräsentation und Seminarbericht

Voraussetzungen:

- dies ist ein *Fünftsemesterkurs*
- Sie sollten den Kurs *Data Analytics Grundlagen* bereits (erfolgreich) absolviert haben
- grundlegende Programmierkenntnisse in R oder Python

Achtung:

Wenn Sie diese **Voraussetzungen nicht erfüllen** und dennoch an dem Kurs teilnehmen möchten, **melden Sie sich bitte** und erklären Sie kurz Ihre Situation.

- von jetzt bis **Donnerstag, 14. Oktober**
 - Gruppenbildung
 - Prioritätenliste festlegen
 - Prioritäten per E-Mail an beide Kursorganisatoren schicken
- wir informieren Sie über die zugeordneten Projekte (vermutlich bis zum 15. Oktober)
- Anmeldung des Kurses beim Prüfungsamt (nachdem die Anmeldefrist abgelaufen ist können Sie sich auch nicht mehr abmelden, nicht abgegebene Projekte zählen dann als nicht bestanden)
- vor der Weihnachtspause (genaues Datum tbd): Zwischenpräsentation
- vor Vorlesungsende: Abschlusspräsentation und Abgabe des Berichts

1. StadionHack Urban Mobility
2. Tankstellendaten
3. VVO-Dashboard
4. VAMOS
5. NFL Big Data Bowl 2022

Alternativ können Sie ein *eigenes Thema vorschlagen*. Es sollte einen Bezug zu Data Analytics (idealerweise im Kontext von Verkehr oder Mobilität) haben oder zu einem unserer Forschungsthemen passen.

Spezifische Aspekte von Data Analytics: Kursorganisation

Projekt 1: #OpenDataCamp:
StadionHack Urban Mobility

Die *Fangemeinschaft Dynamo* bietet das *Open Data Camp (Hackathon)* in mehreren Themenbereichen der urbanen Mobilität an:

- Hackathon mit 5 Challenges (siehe nächste Folien)
- der Hackathon selbst findet am Wochenende (!) vom 22.-24. Oktober im Rudolf-Harbig-Stadium statt (also in etwa einer Woche)

Die *Fangemeinschaft Dynamo* bietet das *Open Data Camp (Hackathon)* in mehreren Themenbereichen der urbanen Mobilität an:

- Hackathon mit 5 Challenges (siehe nächste Folien)
- der Hackathon selbst findet am Wochenende (!) vom 22.-24. Oktober im Rudolf-Harbig-Stadium statt (also in etwa einer Woche)
- Registrierungs-Deadline (von den Hackathon-Organisatoren) ist der **17. Oktober**

Projekt 1: StadionHack Urban Mobility

Die *Fangemeinschaft Dynamo* bietet das *Open Data Camp (Hackathon)* in mehreren Themenbereichen der urbanen Mobilität an:

- Hackathon mit 5 Challenges (siehe nächste Folien)
- der Hackathon selbst findet am Wochenende (!) vom 22.-24. Oktober im Rudolf-Harbig-Stadium statt (also in etwa einer Woche)
- Registrierungs-Deadline (von den Hackathon-Organisatoren) ist der **17. Oktober**

Wenn Sie im Kontext dieses Kurses am Hackathon teilnehmen möchten, **melden Sie sich bitte bei uns beiden via E-Mail vor der Registrierung** und nennen Sie auch die Challenge, an der Sie interessiert sind.

Weitere Details: <https://www.fangemeinschaft-dynamo.de/opendatacamp/>

Besonderer Ablauf:

- die meiste (Programmier-)Arbeit findet am Hackathon-Wochenende statt
- danach:
 - Sie *können* weiter an Ihren Ergebnissen arbeiten
 - Ergebnisse für Präsentation und Bericht aufbereiten
 - Möglichkeit, den Kurs bereits vor der Weihnachtspause zu beenden (die Zwischenpräsentation wird dabei zur Abschlusspräsentation)

Übersicht der Challenges

- Challenge 1: Entspannt zum Ziel
- Challenge 2: Barrierefreiheitsatlas
- Challenge 3: Routing im Stadion
- Challenge 4: History on the Go
- Challenge 5: Besucherstromprognose
- Challenge 6: Wild Card

Spezifische Aspekte von Data Analytics: Kursorganisation

Projekt 2: Tankstellendaten

Projekt 2: Tankstellendaten

- auf den Seiten von [Tankerkönig](https://creativecommons.tankerkoenig.de) (<https://creativecommons.tankerkoenig.de>) gibt es sowohl eine Echtzeit-Benzinpreis-API, sowie historische Benzinpreise zu zahlreichen Tankstellen in ganz Deutschland
- entwickeln Sie ein Shiny-Dashboard mit folgenden Funktionalitäten:
 - interaktive Darstellung der verfügbaren Tankstellen (u.a. für bestimmte Kartenausschnitte, Marken, noch geöffnet, etc.)
 - explorative Datenanalyse der Tankstellenpreise an den verschiedenen Standorten
→ ähnlich zu <https://mehr-tanken.de>, nur besser ;-)
 - Anzeige der schnellsten Strecke(n) von A nach B mit zusätzlicher Anzeige möglicher Tankstellen entlang des Weges (sowie einer Preisprognose für die zu erwartende Ankunftszeit an der Tankstelle)
- zeigen Sie die Funktionalität Ihres Dashboards an verschiedenen Use Cases

Spezifische Aspekte von Data Analytics: Kursorganisation

Projekt 3: VVO Dashboard

- viele Daten des Verkehrsverbund Oberelbe (VVO) stehen per API zur Verfügung, einschließlich der Dresdener Verkehrsbetriebe (DVB)
- umfasst Fahrpläne, aktuelle Verkehrsinformationen, etc.
- Übersicht der APIs: <https://github.com/kiliancoe/vvo> (inoffizielle Sammlung/Dokumentation der Schnittstellen)

Ziel des Projektes:

- Shiny-Dashboard für einen Aspekt der VVO-Daten bauen
- das Dashboard sollte im Ergebnis eine interaktive Analyse/Exploration der Daten erlauben

Spezifische Aspekte von Data Analytics: Kursorganisation

Projekt 4: VAMOS

- Kooperation mit der Professur für Verkehrsleitsysteme und -Prozessautomatisierung
- Analyse der Daten aus dem Verkehrsmanagementsystem VAMOS4
- vorhandene Daten: Verkehrsfluss (insb. Anzahl an Fahrzeugen) im zeitlichen Verlauf der letzten paar Jahre an wichtigen Straßenkreuzungen in Dresden

Ziel des Projektes:

- Visualisierung und Analyse der Verkehrsdaten
- ggf. als interaktives Shiny-Dashboard

Spezifische Aspekte von Data Analytics: Kursorganisation

Projekt 5: NFL Big Data Bowl 2022

Projekt 5: NFL Big Data Bowl 2022

- Kaggle competition basierend auf Daten der weltweit bekanntesten Liga der Sportart American Football
- neben Offense und Defense gibt es beim Football sogenannte Special Teams

Projekt 5: NFL Big Data Bowl 2022

- Kaggle competition basierend auf Daten der weltweit bekanntesten Liga der Sportart American Football
- neben Offense und Defense gibt es beim Football sogenannte Special Teams
- Datenbasis: NFL's Next Gen Stats (NGS) Tracking-Daten aller Special Teams Spielzüge der Jahre 2018-2020 (u.a. Position, Geschwindigkeit, Beschleunigung und Richtung jedes Spielers)
- Ziel: Erzeugung gewinnbringender Informationen anhand der Daten (bspw. Entwicklung neuer Special Teams Metriken, Quantifizierung bestimmter Strategien, oder Ranking der verschiedenen Spieler)

Projekt 5: NFL Big Data Bowl 2022

- Kaggle competition basierend auf Daten der weltweit bekanntesten Liga der Sportart American Football
- neben Offense und Defense gibt es beim Football sogenannte Special Teams
- Datenbasis: NFL's Next Gen Stats (NGS) Tracking-Daten aller Special Teams Spielzüge der Jahre 2018-2020 (u.a. Position, Geschwindigkeit, Beschleunigung und Richtung jedes Spielers)
- Ziel: Erzeugung gewinnbringender Informationen anhand der Daten (bspw. Entwicklung neuer Special Teams Metriken, Quantifizierung bestimmter Strategien, oder Ranking der verschiedenen Spieler)
- Deadline: 6. Januar 2022
- Möglichkeit Preisgeld zu gewinnen

Projekt 5: NFL Big Data Bowl 2022

- Kaggle competition basierend auf Daten der weltweit bekanntesten Liga der Sportart American Football
- neben Offense und Defense gibt es beim Football sogenannte Special Teams
- Datenbasis: NFL's Next Gen Stats (NGS) Tracking-Daten aller Special Teams Spielzüge der Jahre 2018-2020 (u.a. Position, Geschwindigkeit, Beschleunigung und Richtung jedes Spielers)
- Ziel: Erzeugung gewinnbringender Informationen anhand der Daten (bspw. Entwicklung neuer Special Teams Metriken, Quantifizierung bestimmter Strategien, oder Ranking der verschiedenen Spieler)
- Deadline: 6. Januar 2022
- Möglichkeit Preisgeld zu gewinnen
- detaillierte Informationen:
<https://www.kaggle.com/c/nfl-big-data-bowl-2022>

- im Fall des StadionHack-Hackathons, Teilnahme am Hackathon
- Zwischenpräsentation im Dezember
- Abgabe des Programmcodes und Seminarberichts parallel zum Zeitpunkt der Abgabe der Abschlusspräsentationen
- Abschlusspräsentation gegen Vorlesungsende

Fragen und Kommentare?