

To-Do's

Teil 1: Business Understanding % Datenexploration

1. Business Context

- konkretes Business Problem ausformulieren
- konkrete Geschäftsfragen ausformulieren 3-4
- Warum Data Science relevant ausformulieren

2. Explorative Datenanalyse

- 5 verschiedene Datenvizualisierungen
 - StartUps nach Region -> Kreisdiagramm
 - verschiedene Industries -> Säulendiagramm
 - Funding Amount pro Runde mit Farbe/profitabel -> Scatterplot
 - Profitabel mit Revenue -> Boxplot
 - Exit Status mit Founding Year und Dauer -> Line Chart
- Muster, Anomalien, Zusammenhänge feststellen, darstellen und ausformulieren

- potenzielle Biases ausformulieren und darstellen

3. Statistische Analyse

- relevante Statistiken berechnen?
- Korrelationsanalysen durchführen

Teil 2: Datenaufbereitung & Feature Engineering

1. Datenqualität

- Analyse und Behandlung fehlender Werte
- Identifikation und Behandlung von Ausreißer

2. Feature Engineering

- wenn notwendig neue Features, Normalisierung und One-Hot Encoding

3. Dokumentation

- Data Dictionary für alle Features

- Feature 1
- Feature 2
- Feature 3

- Dokumentation der Transformationsschritte

Teil 3: Modellierung & Evaluation

1. Modellauswahl

- Modell begründet auswählen, ausformulieren und implementieren

2. Evaluation

- angemessene Metriken verwenden?
- Diskussion Overfitting/Underfitting?

Teil 4: Kritische Reflexion & Geschäftsempfehlungen

1. Geschäftsempfehlungen

- technische Ergebnisse in Business Sprache übersetzen
- konkrete Empfehlungen ausformulieren
- Diskussion Grenzen der Analyse ausformulieren

2. Weiterführende Fragen

- kritische Fragen, die Analyse aufwirft

- Ausblick für erweiterte Analyse ausformulieren (Schritte..)