

Erstelle ein Jupyter Notebook, erzeuge aus `titanic.csv` ein `DataFrame` u beantworte folgende Fragen:



- **Titel aus dem Namen extrahieren**

Schreibe eine Funktion, die den Titel (z. B. "Mr.", "Miss", "Dr.") aus der `Name` -Spalte extrahiert. Verwende `apply` , um diese Funktion auf die `Name` -Spalte anzuwenden und erstelle eine neue Spalte `Title` . Anschließend verwende `value_counts` , um die Häufigkeit jedes Titels zu bestimmen. Der Name beginnt mit dem Titel. Dieser wird mit einem `.` beendet. (siehe String-Methode `split()`)

- **Überlebenswahrscheinlichkeit nach Titel schätzen**

Verwende die zuvor erstellte `Title` -Spalte und berechne für jede Titelkategorie den durchschnittlichen Überlebensstatus (`Survived`). Nutze `apply` und `groupby` , um die Überlebensrate nach Titel darzustellen (sortiert nach Überlebensrate).

- **Länge der Namen analysieren**

Ersetze die Werte der Spalte `Survived` mit `alive` bzw. `dead` .

Berechne die Anzahl der Zeichen im `Name` jeder Person und erstelle eine neue Spalte `NameLength` . Verwende `apply` und bestimme die durchschnittliche Namenslänge für Überlebende und Nicht-Überlebende.

- **Überlebensrate basierend auf dem Fahrpreis**

Teile die Passagiere in drei Gruppen anhand des Fahrpreises (`Fare`) auf. Überlege, wie du diese Gruppen ungefähr gleichmäßig verteilen kannst. Berechne für jede der drei Gruppen die durchschnittliche Überlebensrate und gib sie aus. (das `DataFrame` kann mit `sort_values(<Spaltenname>)` sortiert werden)

