

Entwicklung von Treibstoffpreisen in Brasilien

Data Preparation

Der Hauptdatensatz, die Treibstoffpreise in Brasilien, wurde zu Beginn einigermaßen umfangreich analysiert und angepasst. So zeigte die erste Ansicht, dass eine Reihe von Features für diese Arbeit nicht zum Einsatz kommen würde – und deswegen entfernt.

Es zeigte sich (leider) auch, dass das wichtige Feature „Mittlerer Verkaufspreis“ in den Originaldaten unbrauchbar ist. Die Einheiten stimmten nicht, zudem fehlten Werte. Doch glücklicherweise gelang es für alle Datensätze, dieses Feature aus den beiden Features „Mindestverkaufspreis“ und „Höchster Verkaufspreis“ neu zu berechnen.

Anschließend wurden aus dem Hauptdatensatz Subsets für die einzelnen Treibstoffarten erstellt.

Im gesamten Prozess erfolgte hier und da Übersetzungsarbeit von Portugiesisch in Deutsch.