Notizen Projektskizze

Gliederung

Thema: „Gebrauchtwagenmarkt Deutschland“ & Datensatz

Datensatz aufbereiten: Autos ohne Wert bei folgenden Variablen:

(name), price, vehicle\_typ, registration\_year, gearbox, power, model, odometer, fuel, brand, ad\_created, postal\_code, last\_seen

Deskriptive Analyse / Marktanalyse

* Markenanteile -> Kuchendiagramm
* Fahrzeugtypen -> Säulendiagramm
* Baujahr -> Kurve / Liniendiagramm

Fragestellungen

1. Welche Trends lassen sich bezüglich Fahrzeugtyp, Getriebeart, Kraftstoff, Marke erkennen?

* Anzahl der jeweiligen Varianten der Variable pro Gesamtanzahl des Jahres, geplottet als Liniendiagramm (wie deskriptive Analyse Baujahre)

1. Welche Gebrauchtwägen (mit welchen Eigenschaften) haben sich am schnellsten verkauft?

* Unterteilung der Gebrauchtwägen in Gruppen: Differenz von „last\_seen“ und „ad\_created“ unterteilt in: < 1 Woche; zwischen 1 Woche und 4 Wochen; > 4 Wochen
* Fahrzeugtypen mit größten Anteilen -> Marken mit größten Anteilen: Clustering / Punktdiagramm von Leistung und Kilometerstand

1. Was ist ein realistischer Preis für mein Wunsch-Fahrzeug?

* Nutzereingabe: Fahrzeugtyp, Marke, Getriebeart, Kraftstoffart } optional?

Leistung, Kilometerstand, Baujahr +/- 3 Jahre

* Auswahl der Daten: Fahrzeugtyp -> Marke -> Getriebeart -> Kraftstoffart
* Nearest Neighbor: Leistung, Kilometerstand, Baujahr +/- 3 Jahre -> Durchschnitt der Preise

1. Karte: In welchen Landkreisen ist die Anzahl der verkauften Gebrauchtwägen am größten? / Wo lohnt es sich am meisten sich als Gebrauchtwagenhändler niederzulassen?