

Krankenversicherungskosten

Aufgabenübersicht:

Du bist bei einer Versicherung angestellt und hast Daten zu verschiedenen Versicherungsnehmern und den Behandlungskosten, die ihnen die Versicherung in Rechnung gestellt hat. Du sollst die Faktoren analysieren, die die Kosten in die Höhe treiben. Das Ziel ist es nämlich, ein Modell zu bauen, das bei einer neuen Behandlung die entsprechenden Kosten automatisch nach einem festen Schema errechnet, sodass die Verwaltung endlich aufhören kann mit Excel herumzuspielen. Am Ende will dein Chef deine Analyse sehen (er liebt Grafiken!) und außerdem wünscht er sich auch ein paar „Predictions“. Er würde also gern ein paar Kosten-Prognosen für ein paar Fälle von deinem Modell erhalten. Im Raum wird übrigens auch noch eine weitere Person aus der Verwaltung sein, die für ethische Standards verantwortlich ist (seit neustem gibt es ein Gesetz, dass mindestens ein Mitarbeiter in größeren Unternehmen in dieser Funktion prüfen soll, ob alles „mit rechten Dingen“ zugeht). Es ist dir überlassen, ob du bei deiner Präsentation also auch diesen Umstand nutzen willst.

Datensatzbeschreibung:

Der Datensatz umfasst folgende Informationen zu Personen: Geschlecht, Body-Mass-Index, Anzahl Kinder, Raucherstatus, Landesregion sowie die in Rechnung gestellten Behandlungskosten.