

Folien für das Modul 'Praxis der Datenanalyse'

FOM – WS17

John Doe III.

Orga

Die Funktion `glm` führt die logistische Regression durch.

```
glm1 <- glm(Aktienkauf ~ Risikobereitschaft,  
            family = binomial("logit"),  
            data = Aktien)
```

Kap sdjlfk

Termin	Thema / Kapitel
1	Organisatorisches
1	Einführung
1	Rahmen
1	Daten einlesen
2	Datenjudo
3	Daten visualisieren
4	Fallstudie (z.B. zu 'movies')
5	Daten modellieren
5	Der p-Wert
6	Lineare Regression - metrisch
7	Lineare Regression - kategorial
8	Fallstudie (z.B. zu 'titanic' und 'affairs')
9	Vertiefung 1: Textmining oder Clusteranalyse
10	Vertiefung 2: Dimensionsreduktion
11	Wiederholung

- ▶ Die Prüfung besteht aus zwei Teilen
 - ▶ einer Klausur (50% der Teilnote)
 - ▶ einer Datenanalyse (50% der Teilnote).

Prüfungsrelevant ist der gesamte Stoff aus dem Skript und dem Unterricht mit **einigen Ausnahmen**

Alle Hinweise zur Prüfung gelten nur insoweit nicht anders vom Dozenten festgelegt.

- ▶ Einen Überblick über die fünf wesentliche Schritte der Datenanalyse gewinnen.
- ▶ R und RStudio installieren können.
- ▶ Einige häufige technische Probleme zu lösen wissen.
- ▶ R-Pakete installieren können.
- ▶ Einige grundlegende R-Funktionalitäten verstehen.
- ▶ Auf die Frage “Was ist Statistik?” eine Antwort geben können.

Alle Datensätze liegen im Ordner `data/`, den Sie vom [Github-Repository](#) herunterladen können.

2. Kap sdjflk R und RStudio installieren

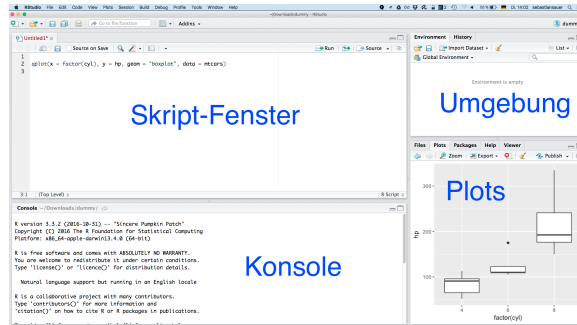


Abbildung 1: RStudio

2. Kap sdjlfk Zeilen zählen mit `n` und `count`

Gruppe A Gruppe B Gruppe C

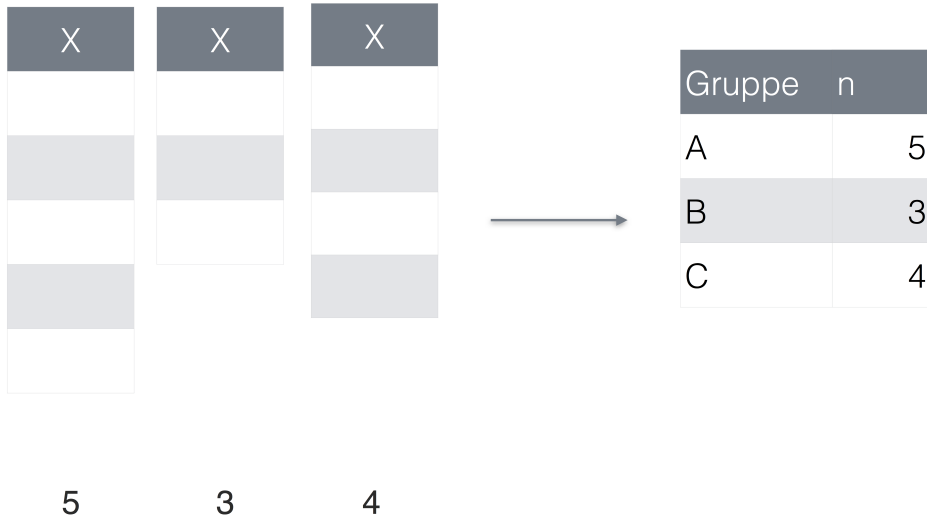


Abbildung 2: Sinnbild für 'count'