Scientific studies

Simpson's paradox

Prof. Dr. Jan Kirenz

Overall gender distribution

Plot Code

```
ggplot(ucbadmit, aes(y = gender, fill = admit)) +
  geom_bar(position = "fill") +
  labs(title = "Admit by gender",
       y = NULL, x = NULL)
```

Gender distribution, by department

Plot Code

```
ggplot(ucbadmit, aes(y = gender, fill = admit)) +
  geom_bar(position = "fill") +
  facet_wrap(. ~ dept) +
  scale_x_continuous(labels = label_percent()) +
  labs(title = "Admissions by gender and department",
       x = NULL, y = NULL, fill = NULL) +
  theme(legend.position = "bottom")
```

Closer look at departments

Output Code

```
ucbadmit %>%
  count(dept, gender, admit) %>%
  group_by(dept, gender) %>%
  mutate(
    n_applied  = sum(n),
    prop_admit = n / n_applied
    ) %>%
  filter(admit == "Admitted") %>%
  rename(n_admitted = n) %>%
  select(-admit) %>%
  print(n = 12)
```