

Data science basics

Recoding data (Case: religion and income)

Prof. Dr. Jan Kirenz

Recode religion

Recode

Plot

```
rel_inc_long <- rel_inc_long %>%  
  mutate(religion = case_when(  
    religion == "Evangelical Protestant" ~ "Ev. Protestant",  
    religion == "Historically Black Protestant" ~ "Hist. Black Protestant",  
    religion == 'Unaffiliated (religious "nones")' ~ "Unaffiliated",  
    TRUE ~ religion  
  ))
```

Reverse religion order

Recode

Plot

```
rel_inc_long <- rel_inc_long %>%  
  mutate(religion = fct_rev(religion))
```

Add income

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col()
```

Fill bars

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill")
```

Change colors

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill") +  
  scale_fill_viridis_d()
```

Change theme

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill") +  
  scale_fill_viridis_d() +  
  theme_minimal()
```

Move legend to the bottom

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill") +  
  scale_fill_viridis_d() +  
  theme_minimal() +  
  theme(legend.position = "bottom")
```


Legend adjustments

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill") +  
  scale_fill_viridis_d() +  
  theme_minimal() +  
  theme(legend.position = "bottom") +  
  guides(fill = guide_legend(nrow = 2, byrow = TRUE))
```

Fix labels

Plot

Code

```
ggplot(rel_inc_long, aes(y = religion, x = frequency, fill = income)) +  
  geom_col(position = "fill") +  
  scale_fill_viridis_d() +  
  theme_minimal() +  
  theme(legend.position = "bottom") +  
  guides(fill = guide_legend(nrow = 2, byrow = TRUE)) +  
  labs(  
    x = "Proportion", y = "",  
    title = "Income distribution by religious group",  
    subtitle = "Source: Pew Research Center, Religious Landscape Study",  
    fill = "Income"  
  )
```