

Simple Regression Models

Introduction

Prof. Dr. Jan Kirenz

... without the measure of uncertainty

Plot Code

```
ggplot(data = pp, aes(x = Width_in, y = Height_in)) +  
  geom_point() +  
  geom_smooth(method = "lm",  
              se = FALSE) +  
  labs(  
    title = "Height vs. width of paintings",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)")
```

... with different cosmetic choices

Plot Code

```
ggplot(data = pp, aes(x = Width_in, y = Height_in)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE,  
              color = "#8E2C90", linetype = "dashed", size = 3) +  
  labs(  
    title = "Height vs. width of paintings",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)"  
)
```

Other smoothing methods: gam

[Plot](#) [Code](#)

```
ggplot(data = pp, aes(x = Width_in, y = Height_in)) +  
  geom_point() +  
  geom_smooth(method = "gam",  
              se = FALSE, color = "#8E2C90") +  
  labs(  
    title = "Height vs. width of paintings",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)"  
)
```

Other smoothing methods: loess

[Plot](#) [Code](#)

```
ggplot(data = pp, aes(x = Width_in, y = Height_in)) +  
  geom_point() +  
  geom_smooth(method = "loess",  
              se = FALSE, color = "#8E2C90") +  
  labs(  
    title = "Height vs. width of paintings",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)"  
)
```

Residuals

Plot Code

```
ht_wt_fit <- linear_reg() %>%
  set_engine("lm") %>%
  fit(Height_in ~ Width_in, data = pp)

ht_wt_fit_tidy <- tidy(ht_wt_fit$fit)
ht_wt_fit_aug <- augment(ht_wt_fit$fit) %>%
  mutate(res_cat = ifelse(.resid > 0, TRUE, FALSE))

ggplot(data = ht_wt_fit_aug) +
  geom_point(aes(x = Width_in, y = Height_in, color = res_cat)) +
  geom_line(aes(x = Width_in, y = .fitted), size = 0.75, color = "#8E2C90") +
  labs(
    title = "Height vs. width of paintings",
    subtitle = "Paris auctions, 1764 - 1780",
    x = "Width (inches)",
    y = "Height (inches)"
  ) +
  guides(color = FALSE) +
  scale_color_manual(values = c("#260b27", "#e6b0e7")) +
  geom_text(aes(x = 0, y = 150), label = "Positive residual", color = "#e6b0e7", hjust = 0, size = 8) +
  geom_text(aes(x = 150, y = 25), label = "Negative residual", color = "#260b27", hjust = 0, size = 8)
```

Landscape paintings

- Landscape painting is the depiction in art of landscapes – natural scenery such as mountains, valleys, trees, rivers, and forests, especially where the main subject is a wide view – with its elements arranged into a coherent composition.¹
 - Landscape paintings tend to be wider than they are long.
- Portrait painting is a genre in painting, where the intent is to depict a human subject.²
 - Portrait paintings tend to be longer than they are wide.

[1] Source: Wikipedia, [Landscape painting](#)

[2] Source: Wikipedia, [Portrait painting](#)

Multiple explanatory variables

Plot Code

```
ggplot(data = pp, aes(x = Width_in, y = Height_in, color = factor(landsALL))) +  
  geom_point(alpha = 0.4) +  
  geom_smooth(method = "lm", se = FALSE) +  
  labs(  
    title = "Height vs. width of paintings, by landscape features",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)",  
    color = "landscape"  
  ) +  
  scale_color_manual(values = c("#E48957", "#071381"))
```

Extending regression lines

Plot Code

```
ggplot(data = pp, aes(x = Width_in, y = Height_in, color = factor(landsALL))) +  
  geom_point(alpha = 0.4) +  
  geom_smooth(method = "lm", se = FALSE,  
              fullrange = TRUE) +  
  labs(  
    title = "Height vs. width of paintings, by landscape features",  
    subtitle = "Paris auctions, 1764 – 1780",  
    x = "Width (inches)",  
    y = "Height (inches)",  
    color = "landscape"  
) +  
  scale_color_manual(values = c("#E48957", "#071381"))
```