

Data Wrangling and Analysis of Experiment Results

Masoud Jasbi

7/8/2019

Data Wrangling

```
qualtrics_data <- read.csv("data/raw/Pilot-Quantifiers-data-trimmed.csv", na.strings = "")

N_dropout <-
  qualtrics_data %>%
  filter(Finished=="FALSE") %>%
  nrow()

N_participants <-
  qualtrics_data %>%
  filter(Finished=="TRUE") %>%
  nrow()
```

Total of 69 took this survey. 13 did not finish the survey and dropped out. The results for 56 are reported.

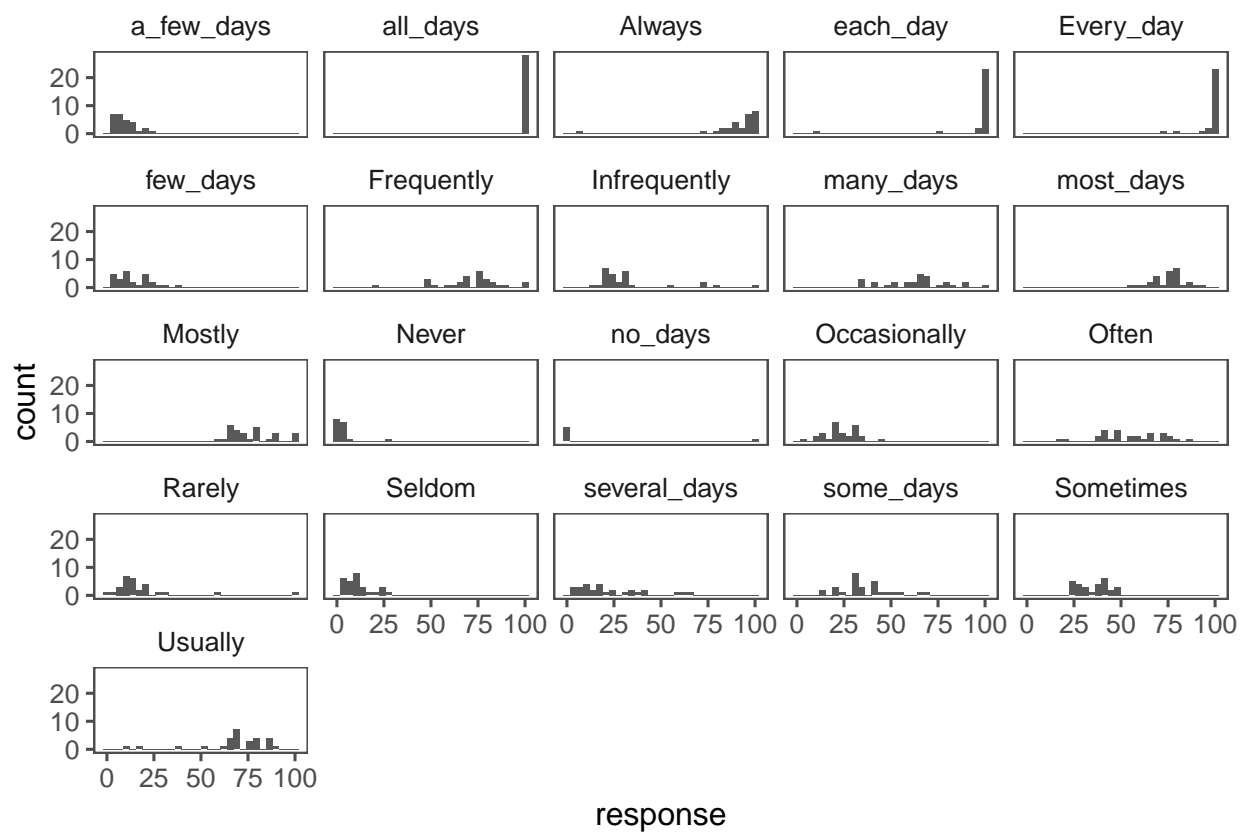
```
tidy_data <-
  qualtrics_data %>%
  filter(Finished=="TRUE") %>%
  mutate(sid=c(1:56)) %>%
  select(-Finished) %>%
  gather(quantifier, response, Every_day:Never) %>%
  drop_na(response) %>%
  mutate(condition = c(rep("nominal", 255), rep("temporal", 297))) %>%
  mutate(response = as.numeric(response))

write_csv(tidy_data, "data/processed/Pilot-Quantifiers-data-processed.csv")
```

Plots

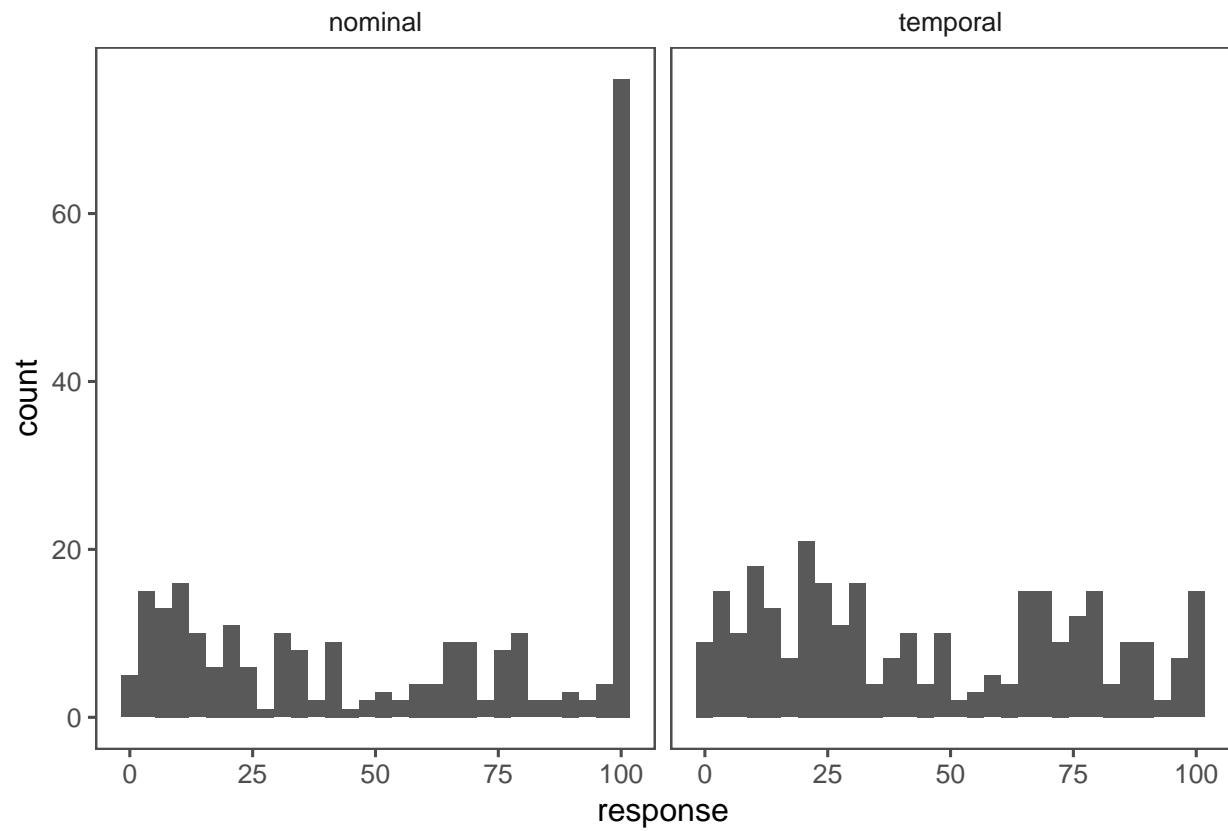
```
tidy_data %>%
  ggplot(aes(response)) +
  geom_histogram() +
  facet_wrap(~quantifier) +
  theme_few()

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

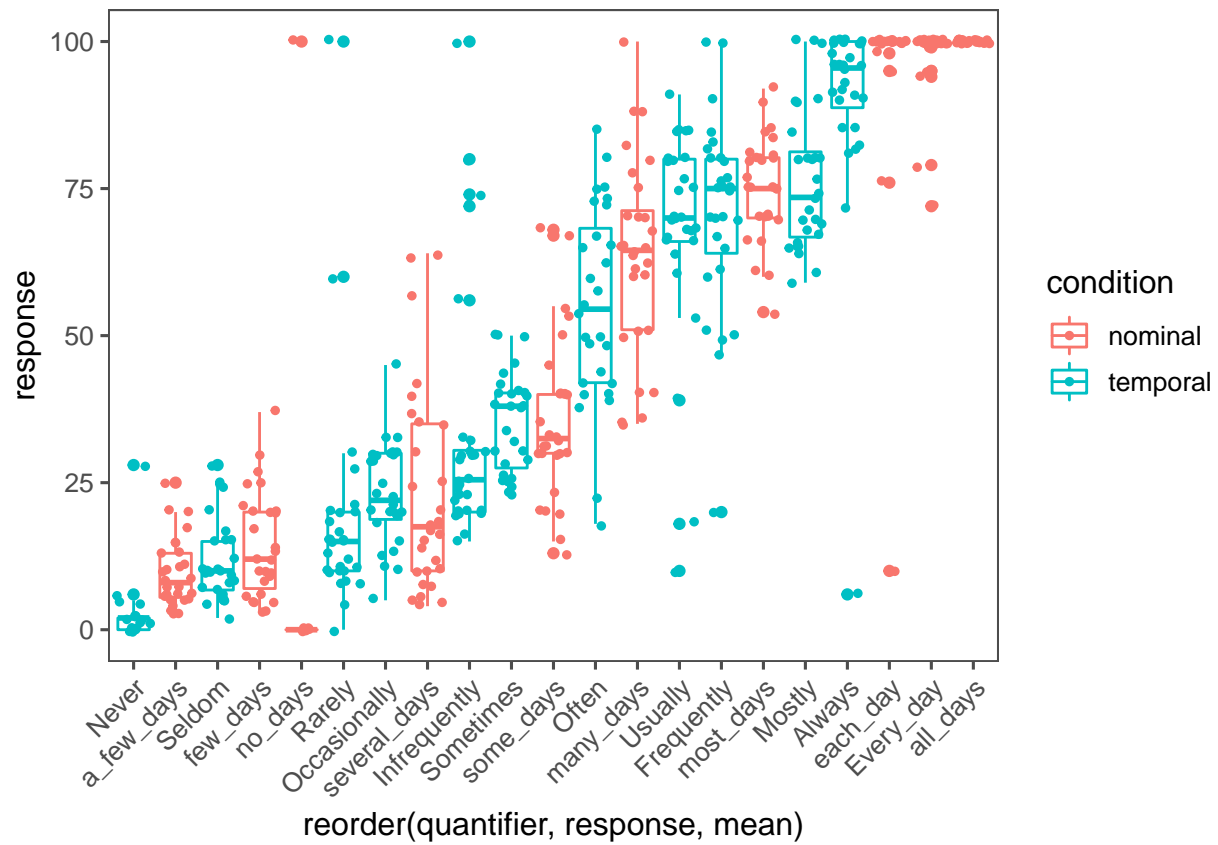


```
tidy_data %>%
  ggplot(aes(response)) +
  geom_histogram() +
  facet_wrap(~condition) +
  theme_few()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
tidy_data %>%
  ggplot(aes(reorder(quantifier, response, mean), response, color=condition), fill=condition) +
  geom_boxplot() +
  geom_jitter(size=1) +
  theme_few() +
  theme(axis.text.x = element_text(angle=45, hjust = 1, vjust = 1))
```



Analysis

K-means Clustering