Practice Tidy Data, Group 6

May 3, 2020

Teresa Burlingame, Julie Sampson,
Josh Olsen,Jake Oetinger, Katie Tallan,
Susan McNerny, David Corbertera

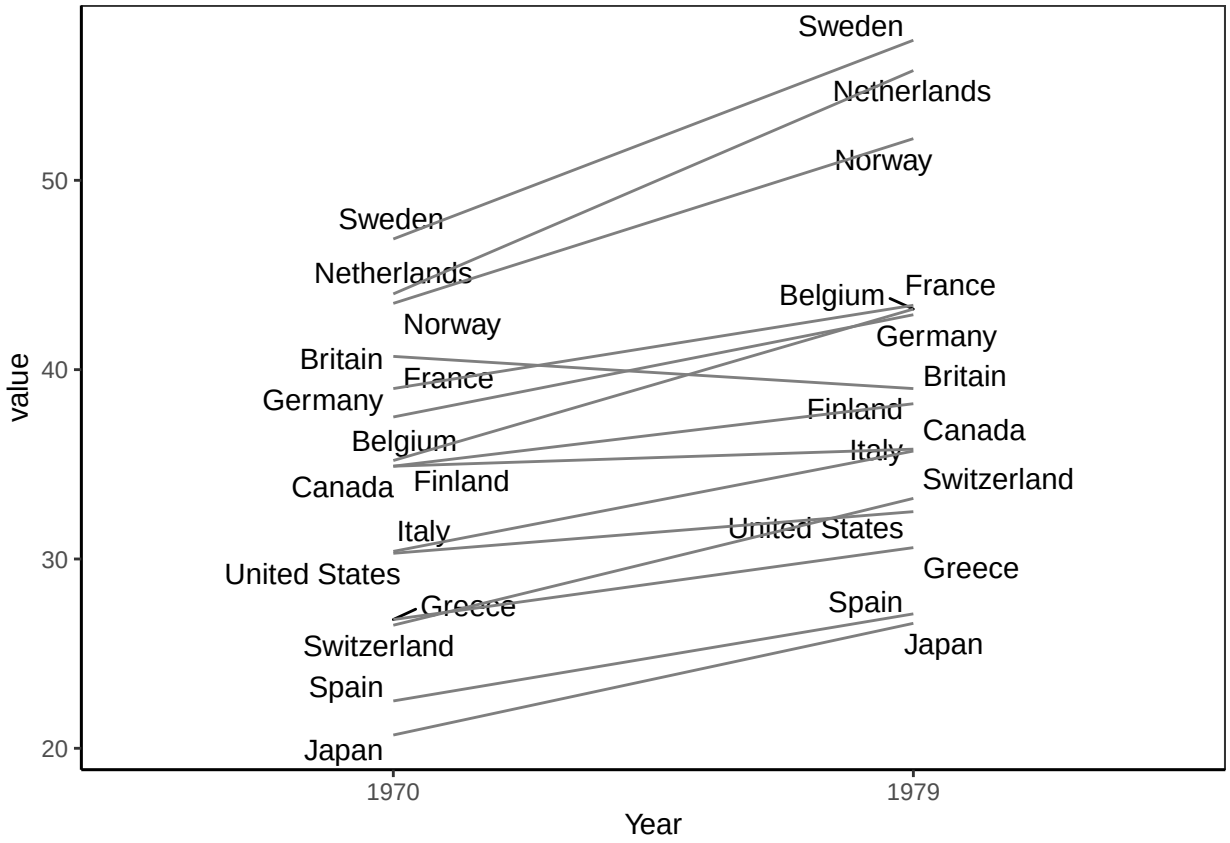
slope Graph - Tufte

```
#install.packages("ggrepel")


# Data Set
tax <- tribble(
  ~ Country,      ~ `1970`, ~ `1979`,
  "Sweden",          46.9,     57.4,
  "Netherlands",     44.0,     55.8,
  "Norway",          43.5,     52.2,
  "Britain",         40.7,     39.0,
  "France",          39.0,     43.4,
  "Germany",         37.5,     42.9,
  "Belgium",         35.2,     43.2,
  "Canada",          34.9,     35.8,
  "Finland",         34.9,     38.2,
  "Italy",           30.4,     35.7,
  "United States",   30.3,     32.5,
  "Greece",          26.8,     30.6,
  "Switzerland",     26.5,     33.2,
  "Spain",           22.5,     27.1,
  "Japan",           20.7,     26.6
)

tax_pivoted <- tax %>% pivot_longer(c("1970","1979"), names_to = "Year")
tax_pivoted
```

```
## # A tibble: 30 x 3
##    Country     Year  value
##    <chr>       <chr> <dbl>
##  1 Sweden      1970   46.9
##  2 Sweden      1979   57.4
##  3 Netherlands 1970   44
##  4 Netherlands 1979   55.8
##  5 Norway      1970   43.5
##  6 Norway      1979   52.2
```

```
##  7 Britain     1970   40.7
##  8 Britain     1979   39
##  9 France      1970   39
## 10 France      1979   43.4
## # ... with 20 more rows
```

```
tax_slope <- ggplot(tax_pivoted, aes(Year, value, label = Country)) +
  geom_text_repel(aes(label=Country)) +
  theme_bw() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))+
  geom_line(aes(group = Country), colour = "grey50")
# p<-ggplot(a) + geom_segment(aes(x=0,xend=months,y=year1,yend=year3),size=.75)
tax_slope
```

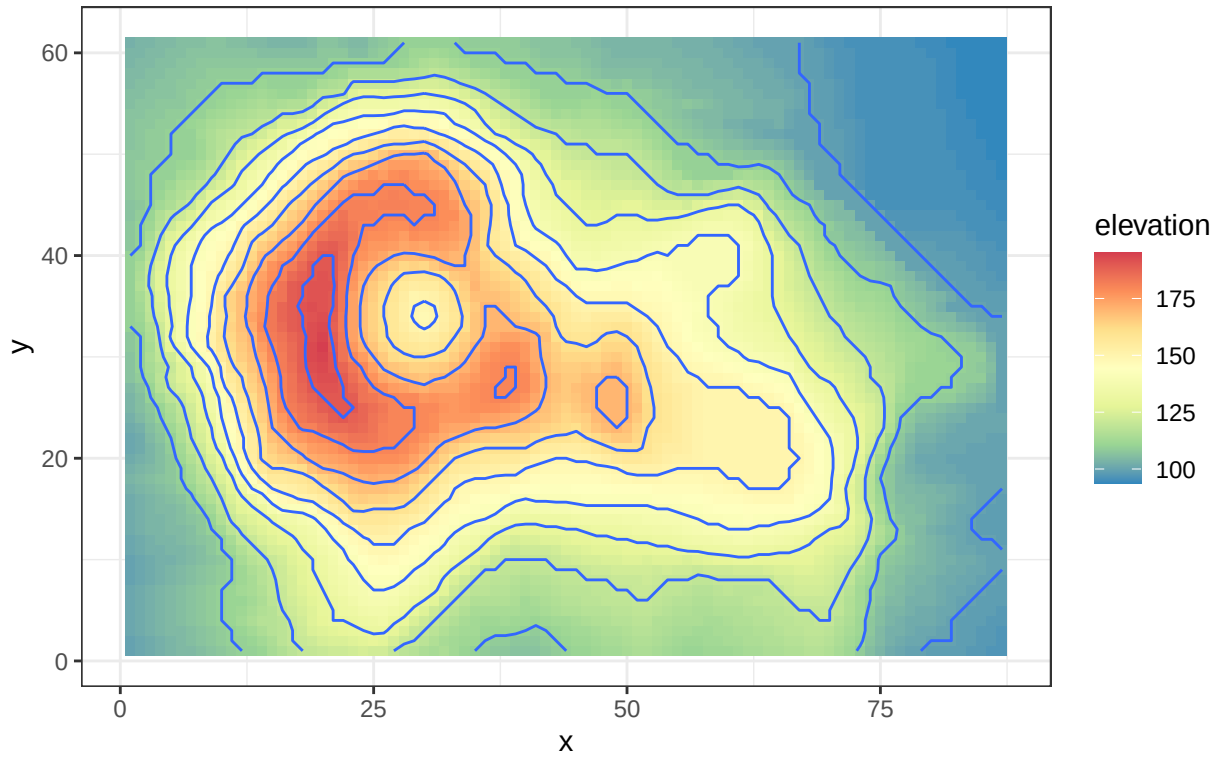



```
# p<-ggplot(a) + geom_segment(aes(x=0,xend=months,y=year1,yend=year3),size=.75)
```

Volcano Heat Map

```
volcano_tbl <- as_tibble(volcano)
colnames(volcano_tbl) <- 1:ncol(volcano)
volcano_tbl$row <- 1:nrow(volcano_tbl)
```

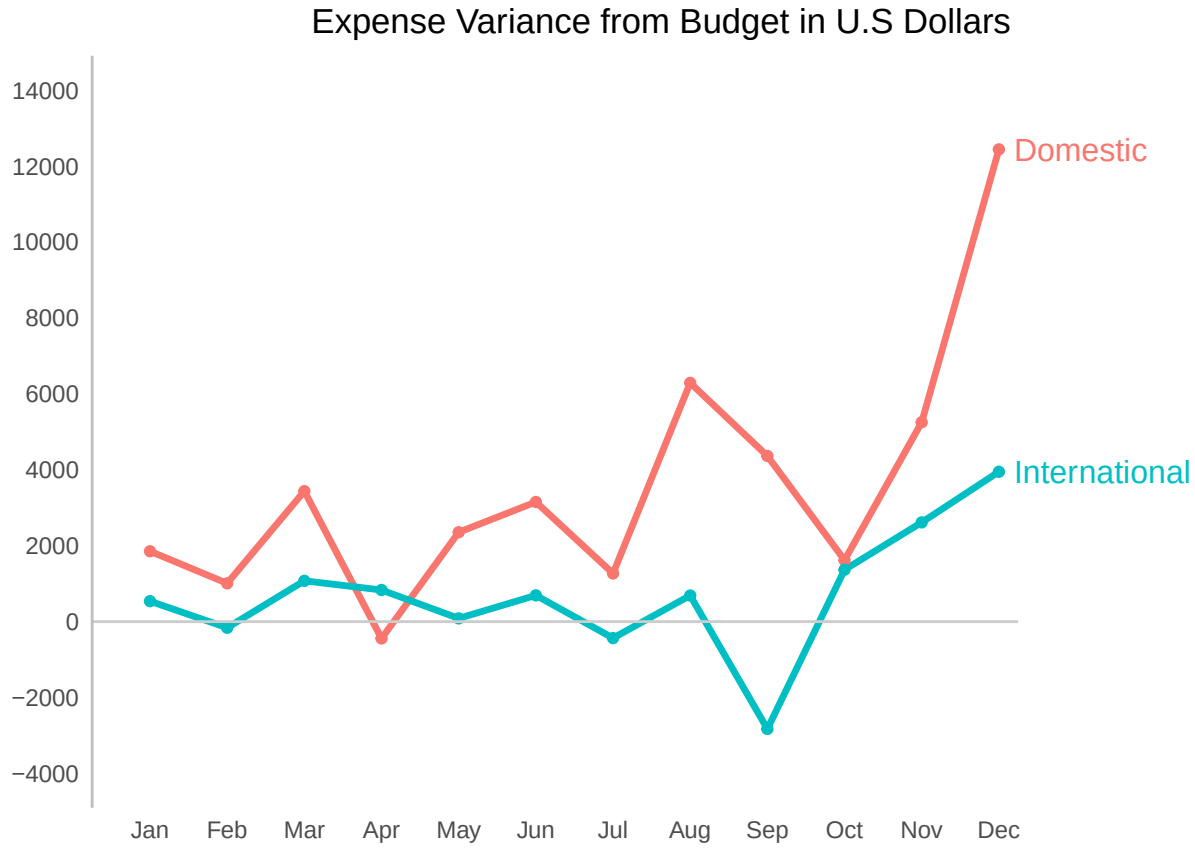
```
heat_map <- volcano_tbl %>% pivot_longer(-contains("row"), names_to = "y",
                                         values_to = "elevation")
names(heat_map) <- c("x", "y", "elevation")
heat_map$y <- as.integer(heat_map$y)
ggplot(heat_map)+
  geom_tile(aes(x,y, fill = elevation))+
  scale_fill_distiller(palette = "Spectral")+
  coord_equal() +
  geom_contour(aes(x = x, y = y, z=elevation)) +
  theme_bw()
```



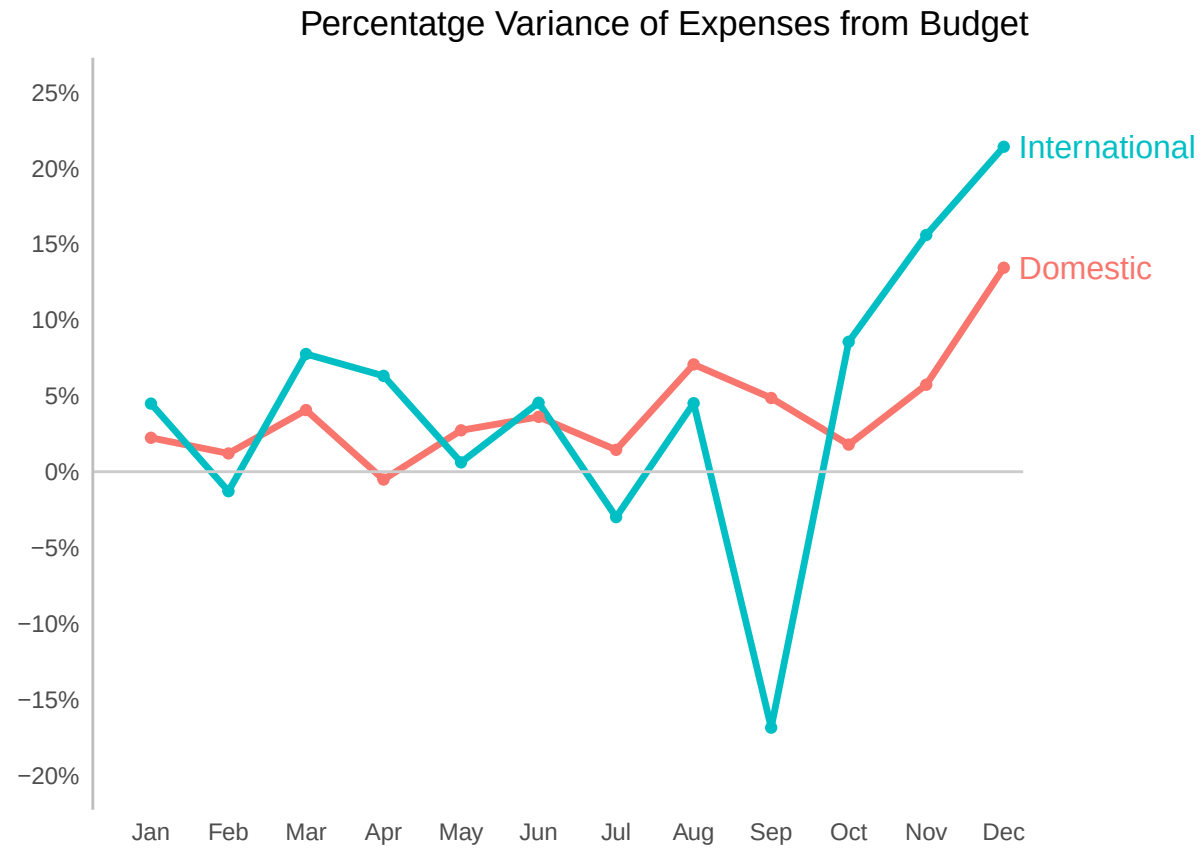
Question 3

```
#Create data set
budget <- tribble(
  ~ Expenses,              ~ Jan, ~ Feb, ~ Mar, ~ Apr, ~ May, ~ Jun, ~ Jul, ~ Aug, ~ Sep, ~ Oct, ~ Nov,
  "Domestic Actual",       84853, 84838, 88103, 85072, 88723, 90384, 89374, 95273, 94239, 92394, 96934,
  "Domestic Budget",       83000, 83830, 84668, 85515, 86370, 87234, 88106, 88987, 89877, 90776, 91684,
  "International Actual", 12538, 12438, 14934, 14033, 13945, 15938, 14086, 15934, 13945, 17338, 19384,
  "International Budget", 12000, 12600, 13860, 13200, 13860, 15246, 14520, 15246, 16771, 15972, 16771,
)

#Tidy it
```

```r
budget2 <- budget %>%
  pivot_longer(cols = -1,
               names_to = "Month",
               values_to = "Amount") %>%
  separate(Expenses, into = c("Type", "Spending")) %>%
  pivot_wider(names_from = Spending,
              values_from = Amount)

#Reorder factors so they plot sequentially
budget2$Month <- budget2$Month %>%
  factor() %>%
  fct_relevel(c("Jan","Feb","Mar","Apr","May","Jun",
                "Jul","Aug","Sep","Oct","Nov","Dec"))

#Figure 9.8
# ggplot(budget2) +
#   geom_line(aes(x = Month, y = Actual-Budget, color = Type, group=Type))

axisx2<-data.frame(Mon=c(0.25:13),y=0)

ggplot(budget2,aes(x = Month, y = Actual-Budget, color = Type, group=Type)) +
  geom_line(size=1.2)+
  geom_point()+
  geom_line(data=axisx2, aes(x=Mon, y=y, group=y), color="grey80")+
  geom_dl(aes(label = Type), method = list(dl.trans(x = x + .2), "last.points"))+
  ggtitle("Expense Variance from Budget in U.S Dollars")+
  theme_classic()+
  theme(plot.title = element_text(hjust = 0.5))+
  theme(legend.position = "none")+
  theme(axis.title = element_blank())+
  theme(axis.line = element_line(color = "grey"))+
  theme(axis.line.x = element_blank())+
  # xlim(c(1,15))+
  expand_limits(x= c(1, 15))+
  expand_limits(y= c(-4000, 14000))+
  theme(axis.ticks=element_blank())+
  #geom_hline(yintercept=0, linetype="dashed", color = "grey")+

  scale_y_continuous(breaks=c(seq(-4000, 14000, by=2000)))
```




```
#Figure 9.9
# ggplot(budget2) +
#   geom_line(aes(x = Month, y = ((Actual-Budget)/Budget)*100, color = Type, group=Type))

ggplot(budget2,aes(x = Month, y = ((Actual-Budget)/Budget)*100, color = Type, group=Type)) +
  geom_line(size=1.2)+
  geom_point()+
  geom_line(data=axisx2, aes(x=Mon, y=y, group=y), color="grey80")+
  geom_dl(aes(label = Type), method = list(dl.trans(x = x + .2), "last.points"))+
  ggtitle("Percentatge Variance of Expenses from Budget")+
  theme_classic()+
  theme(plot.title = element_text(hjust = 0.5))+
  theme(legend.position = "none")+
  theme(axis.title = element_blank())+
  theme(axis.line = element_line(color = "grey"))+
  theme(axis.line.x = element_blank())+
  # xlim(c(1,15))+
  expand_limits(x= c(1, 15))+
  expand_limits(y= c(-20, 25))+
  theme(axis.ticks=element_blank())+
  #geom_hline(yintercept=0, linetype="dashed", color = "grey")+

  scale_y_continuous(breaks=c(seq(-20, 25, by=5)), labels = function(x) paste0(x,"%"))
```


Percentatge Variance of Expenses from Budget
25%
20%
15%
10%
5%
0%
−5%
−10%
−15%
−20%
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
International
Domestic

“‘