Data Analysis for Profit from Policing : Allocation of Civil Asset Forfeiture Funds

Yuhang Li

7/10/2023

```r
#load required packages
library(tidyverse)
library(knitr)
library(scales)
library(huxtable)
```

```r
#load cleaned data sets
national_rev <- read.csv("forfeit_cleaned/national_rev_cleaned.csv")
national_exp <- read.csv("forfeit_cleaned/national_exp_cleaned.csv")
IL_cleaned <- read_csv("forfeit_cleaned/IL_cleaned.csv")
```

```
## Rows: 98153 Columns: 12
## -- Column specification -------------------------------------------------
## Delimiter: ","
## chr (4): agency, state_share, prop_type, exp_type
## dbl (7): id, revenue, nbr_forfeit, agency_share, year, exp_amt, exp_proportion
## lgl (1): procd_type
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
#check data sets
head(national_rev)
```

state	revenue_id	year	unit_type	prop_type	rev	procd_type	conviction	case_id
AR	951	2018	OTHER					0001-18-18E-18
AR	952	2018	OTHER					0002-18-18E-18
AR	953	2018	OTHER					0003-18-8S-319
AR	954	2018	OTHER					0004-18-12-85
AR	955	2018	OTHER					0005-18-12-85
AR	956	2018	OTHER					0006-18-20-141

```r
head(national_exp)
```

```r
head(IL_cleaned)
```

#Part 1: National level data ## Plot number of civil asset forfeiture cases per year

expense_id	year	exp_type	exp_amt	state
121	2017	court costs	10941	FL
122	2017	outside services	1572	FL
123	2017	court costs	2450	FL
124	2017	outside services	610	FL
125	2018	outside services	668	FL
126	2017	travel and training	500	FL

|	revenue	nbr_forfeit	state_share	agency_share	prop_type	procd_type	year	exp
MENT	7.61e+06	5.41e+03			currency		2e+03	
DEPARTMENT	196	1			currency		2e+03	
PARTMENT	3.67e+03	3			currency		2e+03	
PARTMENT	6.7e+03	1			vehicles		2e+03	
TMENT	2.16e+03	1			currency		2e+03	
ARTMENT	2.17e+03	1			currency		2e+03	

```
#plot number of civil asset forfeiture cases over years using line graph
forfeit_nb_years <- national_rev %>%
  ggplot(mapping = aes(x = year)) +
  geom_freqpoly(color = "#144a74") +
  labs(title = "Number of civil asset forfeiture cases over time",
       x = "Year",
       y = "Number of civil asset forfeiture cases") +
  theme_classic()
forfeit_nb_years
```

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

Number of civil asset forfeiture cases over time
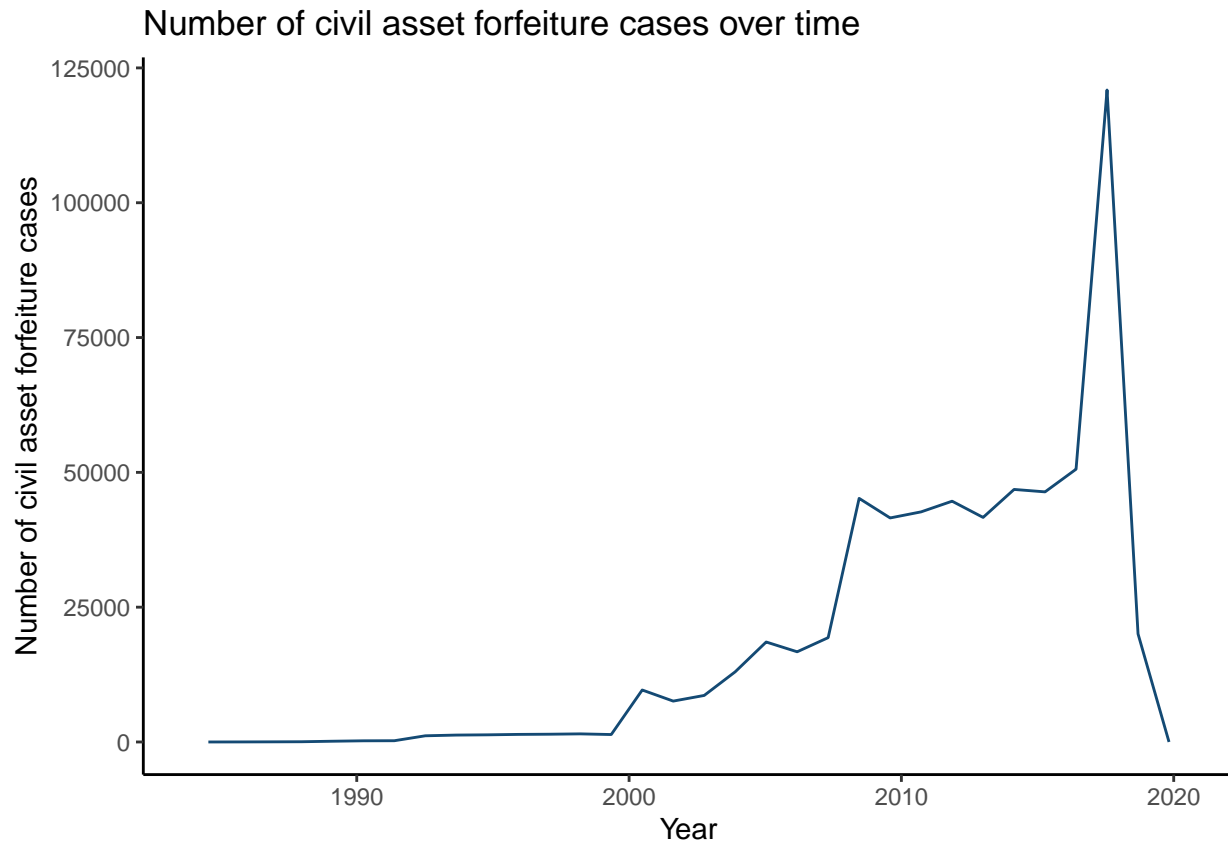
Analyze the trend of average revenue from civil asset forfeiture fund from 1986 to 2019

```
#plot a bar graph of average revenue over years
avg_rev_year <- national_rev %>%
  group_by(year) %>%
  summarize(avg_rev = mean(rev, na.rm = TRUE)) %>%
  ggplot(mapping = aes(x = year, y = avg_rev)) +
  geom_bar(stat = "identity", fill = "#144a74") +
  labs(title = "Average revenue from civil asset forfeiture (1986 - 2019)",
      x = "Year",
      y = "Average revenue in U.S. dollars") +
  theme_classic()
avg_rev_year
```

Average revenue from civil asset forfeiture (1986 – 2019)
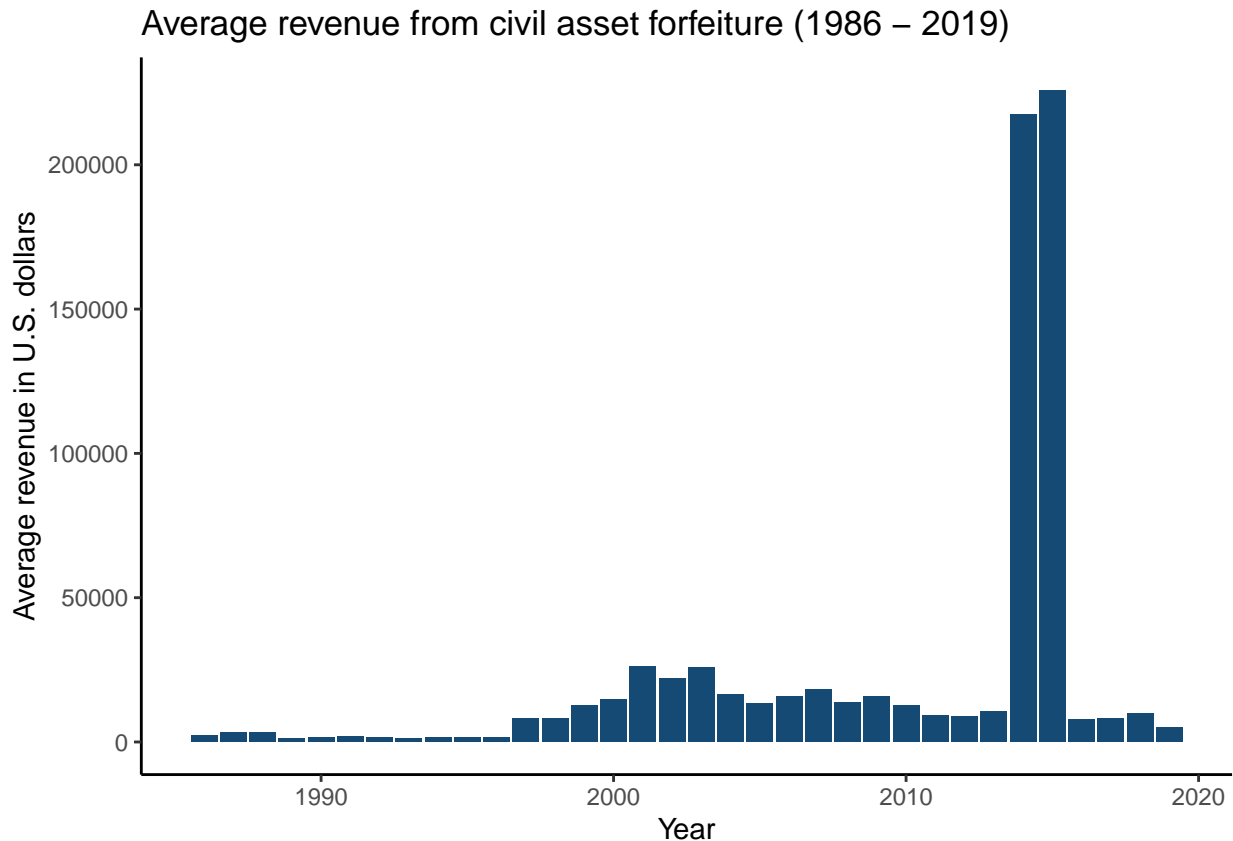
Analyze the proportion of total expenditure for each expenditure type

```r
#write a function to find proportion of total expenditure for each expenditure type
exp_prop <- function(dataset) {
  exp_prop_calc <- dataset %>%
    #filter by positive expenditure values
    filter(exp_type != "NA" & exp_amt > 0) %>%
    group_by(exp_type) %>%
    summarize(total_exp = sum(exp_amt)) %>%
    #calculate proportion of total expenditure
    mutate(exp_pct = total_exp / sum(total_exp))
  return(exp_prop_calc)
}

#find proportion of national expenditure
national_exp_type <- exp_prop(national_exp)

#plot proportion of total expenditure by type of expenditure
exp_type_bar <- national_exp_type %>%
  ggplot(aes(x = reorder(exp_type, exp_pct), y = exp_pct, fill = exp_type)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_brewer(palette = "Set3") +
  labs(title = "Proportion of forfeiture expenditure by expenditure type",
       x = "Proportion of total expenditure",
       y = "Expenditure Type",
```

```
        fill = "Expenditure Type") +
  theme_classic()
exp_type_bar
```

Proportion of forfeiture expenditure by expenditure type
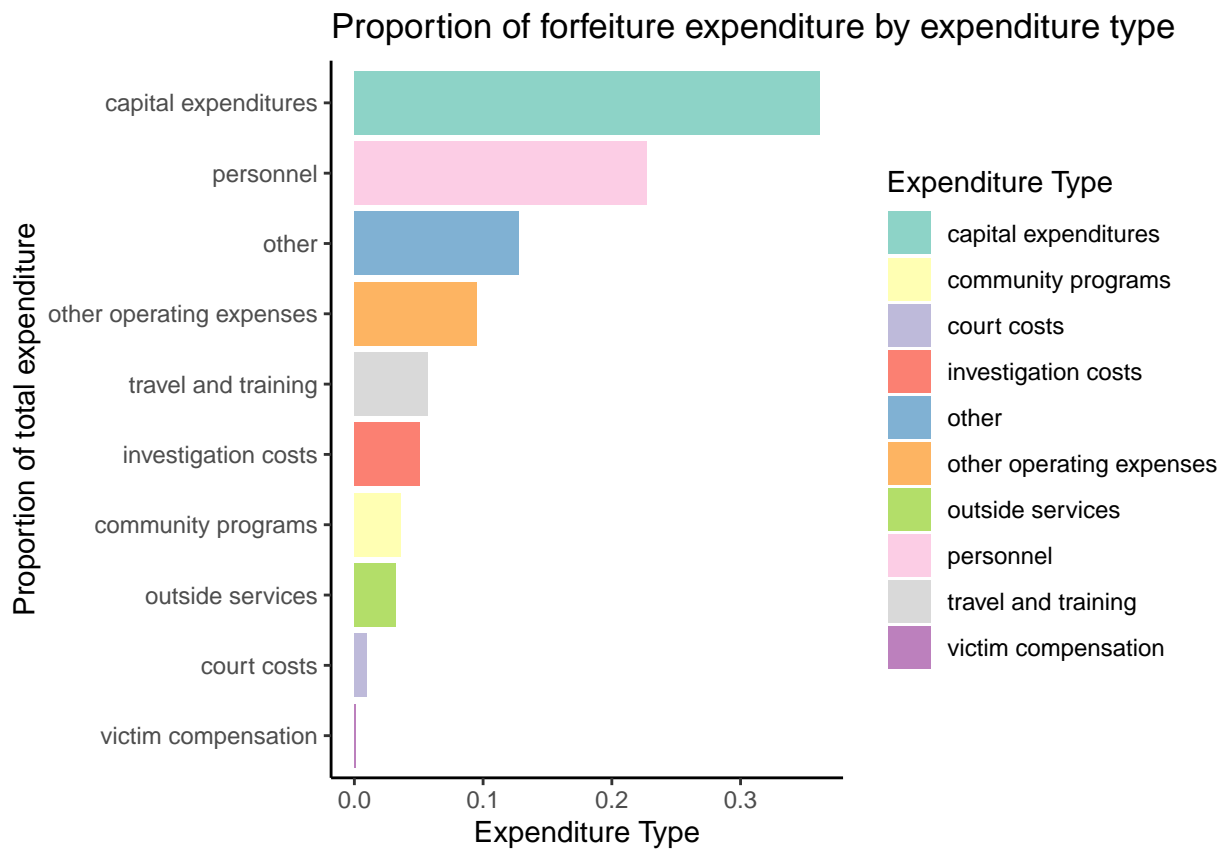
Find the percentage of civil asset forfeitures involving individuals convicted with crime, grouped by property type

```
#find proportion of cases with positive convicted outcome by property type
percent_convicted <- national_rev %>%
  group_by(prop_type) %>%
  summarize(percent_con = mean(conviction, na.rm = TRUE)) %>%
  filter(prop_type != "NA") %>%
  arrange(desc(percent_con))

#format as table
percent_convicted %>%
  #convert to percentage
  mutate(percent_con = percent(percent_con, accuracy = 0.01)) %>%
  kable(caption = "Percentage of civil asset forfeitures with convicted outcomes",
        col.names = c("Property type",
                      "Percent convicted after seizure")
  )
```

Table 1: Percentage of civil asset forfeitures with convicted outcomes

Property type	Percent convicted after seizure
other	17.84%
real property	17.07%
currency	5.19%
vehicles	4.95%

##Find the top 10 and bottom 10 states that had the most and least civil asset forfeitures cases

```r
#find number of cases per state
cases_per_state <- national_rev %>%
  group_by(state) %>%
  summarize(count = n()) %>%
  filter(count > 60) %>%
  arrange(desc(count))

#find top 10 states with most cases
top_10 <- cases_per_state %>%
  head(10)

#find bottom 10 states with least cases
bottom_10 <- cases_per_state %>%
  tail(10)

#combine the two groups
combined_states <- rbind(top_10, bottom_10)

#differentiate the two groups
combined_states$group <- if_else(combined_states$state %in%
                                   top_10$state,
                                 "Top 10",
                                 "Bottom 10")

#plot bargraph
forfeit_rank_states <- combined_states %>%
  ggplot(mapping = aes(x = reorder(state, count), y = count, fill = group)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_manual(values = c("Top 10" = "#599e94",
                               "Bottom 10" = "#d2980d")) +
  labs(title = "Civil asset forfeiture cases 1986 - 2019",
       subtitle = "States that had most and least number of cases",
       x = "Number of civil asset forfeiture cases",
       y = "State",
       fill = "Group") +
  theme_classic()

#print bargraph
forfeit_rank_states
```
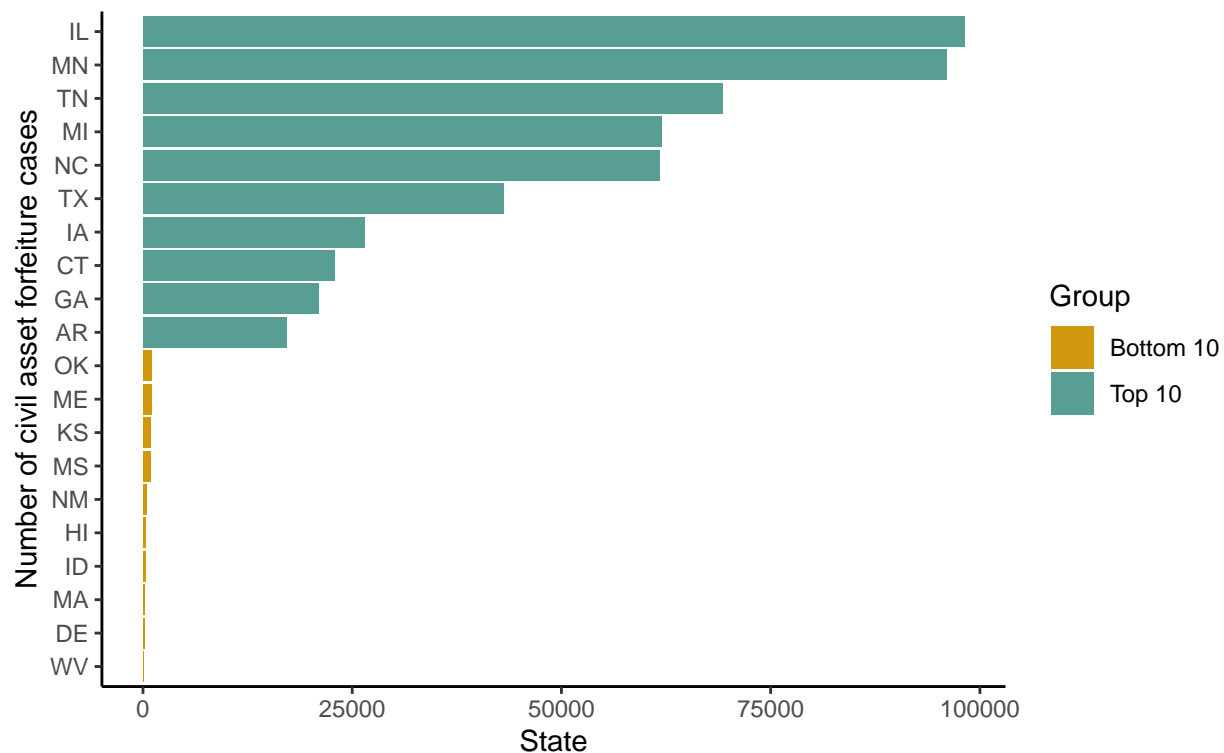
Civil asset forfeiture cases 1986 – 2019

States that had most and least number of cases

#Part 2 : Illinois Data
Summary statistics

```r
#select column names for which summary statistics will be calculated
selected_cols <- c("revenue", "exp_amt", "agency_share")

#create for loop to calculate summary statistics
summary_stats <- list()
for (col in selected_cols) {
  column_summary <- summary(IL_cleaned[[col]], na.rm = TRUE)
  summary_stats[[col]] <- column_summary
}

#format the three sets of summary statistics into a table
summary_table <- hux(Variable = c("revenue", "expenditure", "agency share"),
                     Min = c(0, 0, 0),
                     Q1 = c(278, 0, 64.2),
                     Median = c(714, 0, 175),
                     Q3 = c(2094, 0, 538.4),
                     Max = c(17396274, 3189760, 86450)) %>%
  set_all_padding(4) %>%
  set_outer_padding(0) %>%
  set_number_format(0) %>%
  set_bold(row = 1, col = everywhere) %>%
  set_bottom_border(row = 1, col = everywhere) %>%
  set_width(0.4) %>%
  set_caption("Summary statistics for revenue, expenditure, and agency share")
```

```
#print table
summary_table
```

Table 2: Summary statistics for revenue,
expenditure, and agency share

| Variable	Min	Q1	Median	Q3	Max
revenue	0	278	714	2094	17396274
expenditure	0	0	0	0	3189760
agency share	0	64	175	538	86450

Trend of IL average revenue per year from 2000 to 2019

```
#find average revenue per year
IL_avg <- IL_cleaned %>%
  #remove outlier
  filter(year != "2009") %>%
  group_by(year) %>%
  summarize(avg_rev = mean(revenue, na.rm = TRUE))

#plot a line graph of average revenue over the years in data
IL_avg_rev <- IL_avg %>%
  ggplot(mapping = aes(x = year, y = avg_rev)) +
  geom_freqpoly(stat = "identity", color = "#144a74") +
  labs(title = "Average revenue from IL civil asset forfeiture (2000 - 2019)",
       x = "Year",
       y = "Average revenue in U.S. dollars") +
  theme_classic()

IL_avg_rev
```

Average revenue from IL civil asset forfeiture (2000 – 2019)
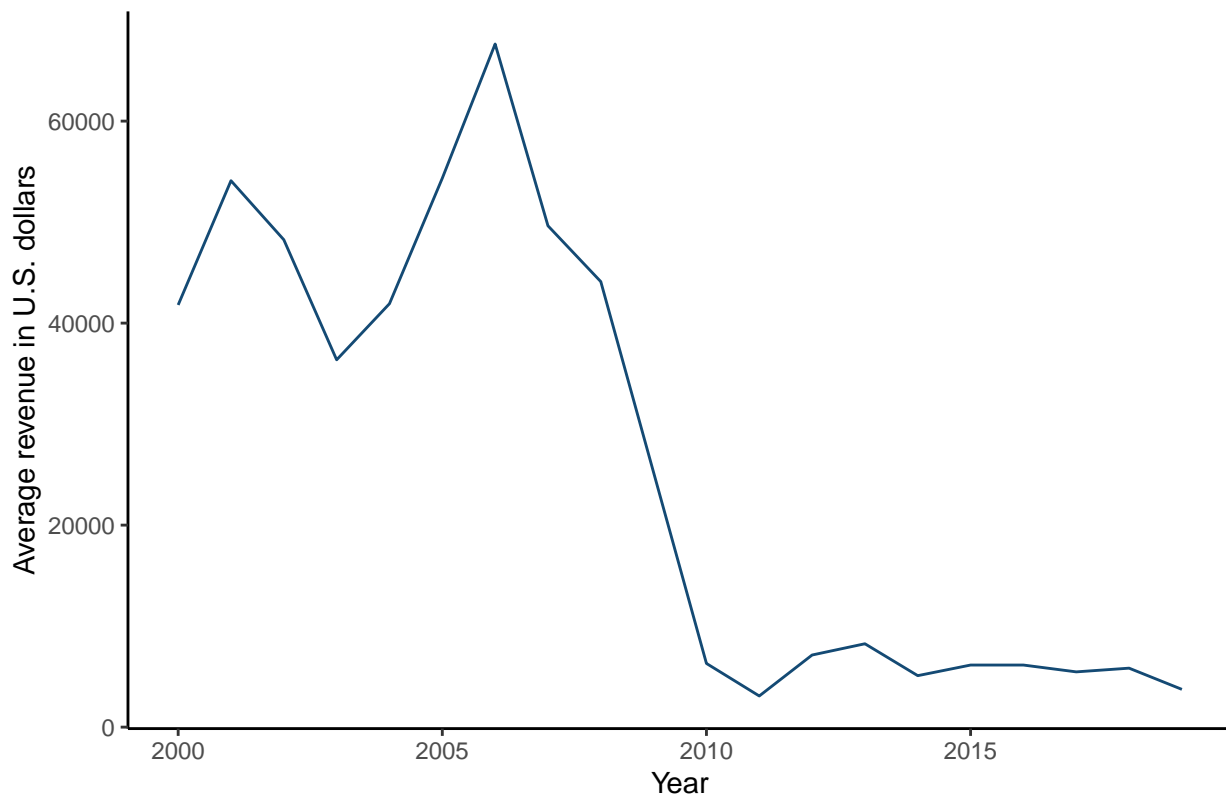
Boxplot of average revenue by property type

```
#plot a boxplot of forfeiture revenue by property type
IL_rev_property <- IL_cleaned %>%
  filter(prop_type != "NA" & revenue < 5000) %>%
  ggplot(mapping = aes(x = prop_type, y = revenue)) +
  geom_boxplot() +
  labs(title = "Revenue by type of property",
       x = "Property type",
       y = "Revenue") +
  coord_flip() +
  theme_classic()
IL_rev_property
```

Revenue by type of property
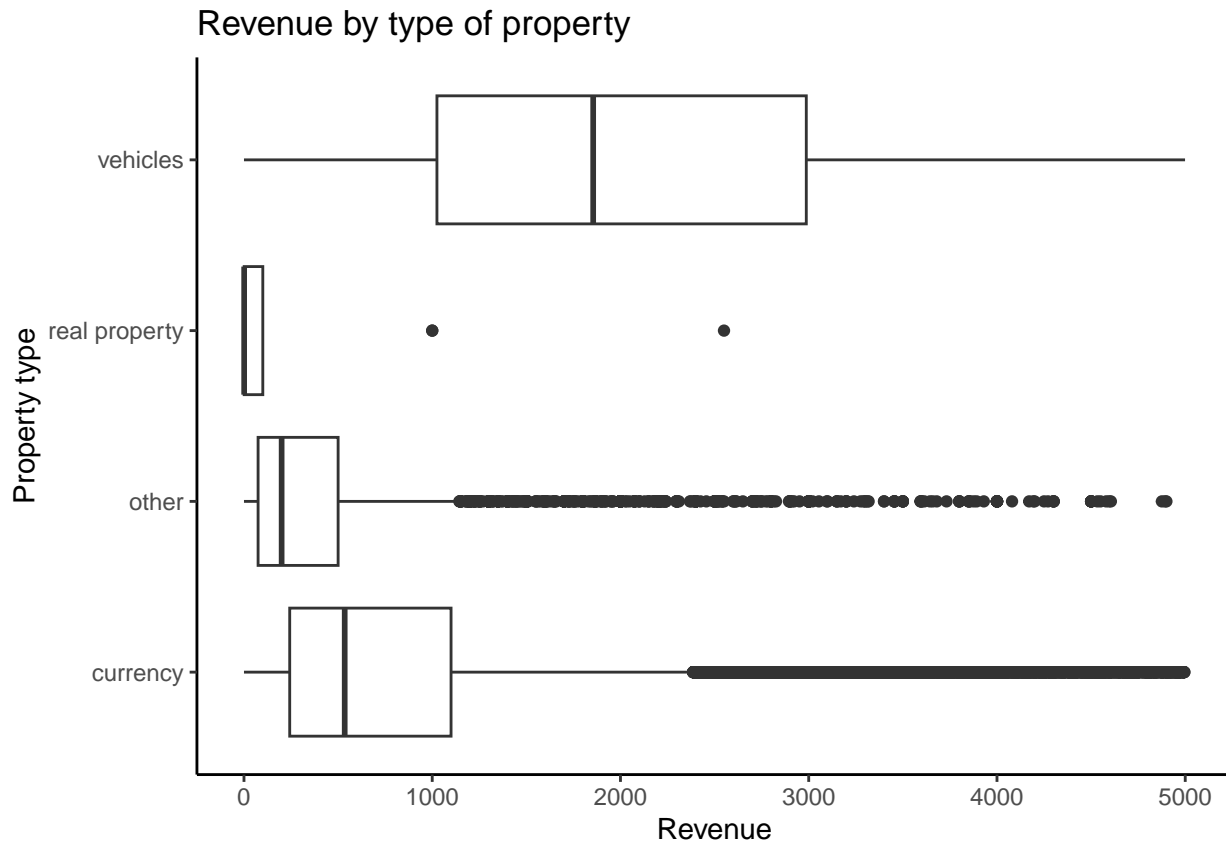
Since distribution of variables are highly skewed, analysis of expenditure is better done by calculating the proportion of total expenditure of each type of expense and visualizing the results in a pie chart.

Pie chart to analyze proportion of total expenditure by expenditure type

'

```
#use previously created function to compute proportion of total expenditure
IL_exp_type <- exp_prop(IL_cleaned)

#create a pie chart to show the proportion of total expenditure occupied by each
#expenditure type
IL_exp_type_pie <- IL_exp_type %>%
  ggplot(aes(x = "", y = exp_pct, fill = exp_type)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar("y") +
  scale_fill_brewer(palette = "Set2") +
  theme_classic() +
  labs(title = "Proportion of total expenditure by type",
       x = NULL,
       y = NULL,
       fill = "Type of Expenditure")
IL_exp_type_pie
```

Proportion of total expenditure by type