

Rose's Work for Final Report

Rose Evard

2023-04-27

```
# make kable table with consistent formatting
make_table <- function(..., title = "", col_names = c("")) {
  df <- as.data.frame(...)
  title <- paste0("<center><span style = 'font-size:150%;color:black'><b>",
                  title,
                  "</span></b><center>")
  as_tibble(df) %>%
  kbl(caption = title,
      col.names = col_names) %>%
  kable_material() %>%
  row_spec(row=0, background = "#43494C", color = "white", bold = TRUE)
}
```

```
#####
# CREATE BIG PALETTE WITH MANY COLORS
#####
pal1 <- viridis_pal(option="mako", end = .9)(12)
pal2 <- viridis_pal(option="rocket", end = .8)(12)
pal3 <- viridis_pal(option="magma", end = .9)(12)
pal4 <- viridis_pal(option="inferno", end = .8)(12)
set.seed(123)
pal <- c(pal1, pal2, pal3, pal4)
indexes <- sample.int(length(pal), replace=FALSE)
pal <- pal[indexes]
```

```
## Loading in data
endowment_data <- read_rds(here("data", "endowments_by_most_recent_filings.RDS"))
names <- readRDS(here("data", "companies.RDS"))
names <- names %>%
  mutate(organization_name = ifelse(
    organization_name == "Ballet Hispanico",
    "Ballet Hispánico",
    organization_name))
```

Calculating Percent Change

Including Investments

(End Year Balance - Beginning Year Balance) / Beginning Year Balance * 100

If EYB is larger, positive result. Meaning there was a INCREASE in total funds.
 If BYB is larger, negative result. Meaning a DECREASE in total funds.
 If result is above 100, the fund was at least DOUBLED.

```
#### USE THIS
## (EYB - BYB)/BYB * 100
## Rose has notes on her choice for this calculation
pct_change_ds <- endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  mutate(change = EndYearBalanceAmt - BeginningYearBalanceAmt,
         pct_change = change/BeginningYearBalanceAmt * 100) %>%
  arrange(desc(pct_change)) %>%
  left_join(names, by = "EIN")
```

REMOVING Investments

$$\frac{(\text{End Year Balance} - \text{Investment Earnings or Losses Amount} - \text{Beginning Year Balance})}{\text{Beginning Year Balance}} * 100$$

Positive result: INCREASE in total funds.
 Negative result: DECREASE in total funds.
 If result is above 100, the fund was at least DOUBLED.

```
#### USE THIS
## (EYB - INV - BYB)/BYB * 100
## Rose has notes on her choice for this calculation
pct_change_inv_ds <- endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  mutate(InvestmentEarningsOrLossesAmt = ifelse(is.na(InvestmentEarningsOrLossesAmt), 0, InvestmentEarningsOrLossesAmt),
         change = EndYearBalanceAmt - InvestmentEarningsOrLossesAmt - BeginningYearBalanceAmt,
         pct_change = change/BeginningYearBalanceAmt * 100) %>%
  arrange(desc(pct_change)) %>%
  left_join(names, by = "EIN")
```

Summarizing Information

With investments

```
# Basic summary stats
pct_change_ds %>%
  filter(!is.na(pct_change) & pct_change != Inf) %>%
  summarize(avg_pct_change = mean(pct_change),
            median_pct_change = median(pct_change),
            sd_pct_change = sd(pct_change))
```

```
## # A tibble: 1 x 3
##   avg_pct_change median_pct_change sd_pct_change
##   <dbl>           <dbl>           <dbl>
## 1      30.6         2.07           201.
```

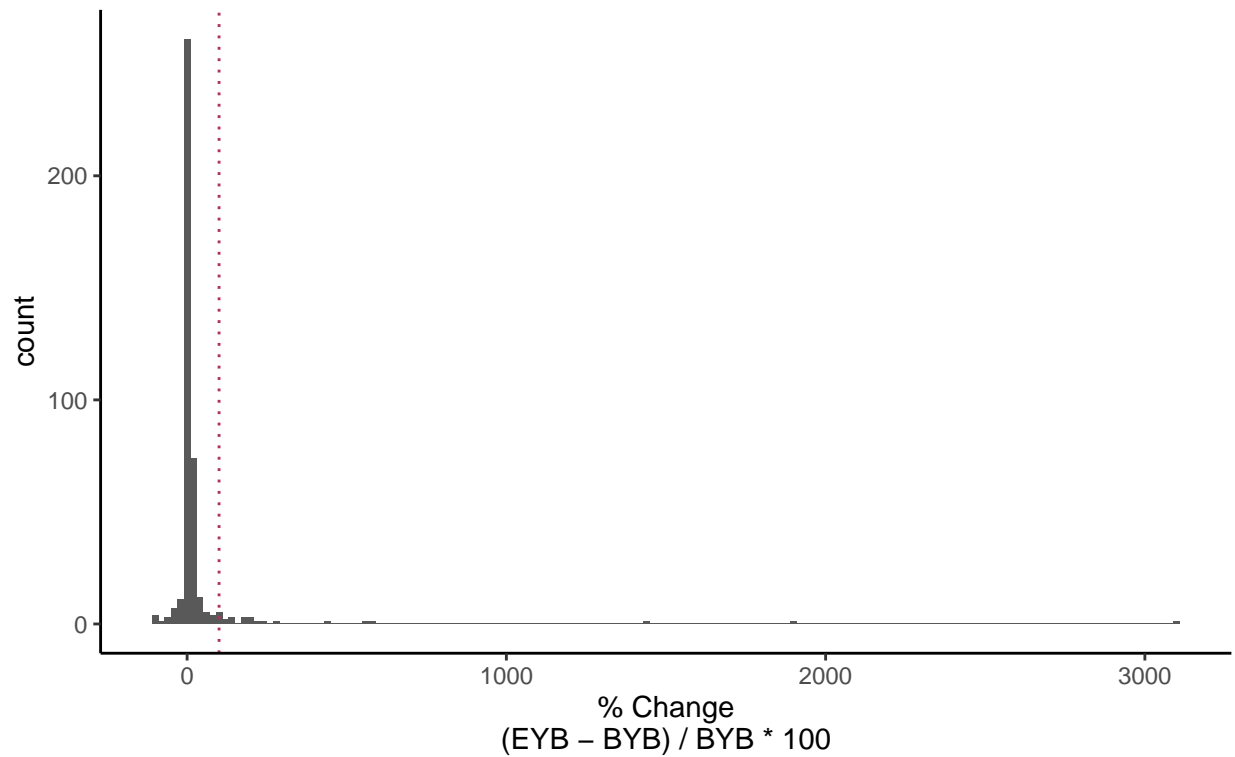
```
pct_change_ds %>%
  filter(!is.na(pct_change) & pct_change != Inf) %>%
  group_by(EIN) %>%
  summarize(avg_pct_change = mean(pct_change),
            median_pct_change = median(pct_change),
            sd_pct_change = sd(pct_change))
```

```
## # A tibble: 46 x 4
##   EIN      avg_pct_change median_pct_change sd_pct_change
##   <chr>          <dbl>          <dbl>          <dbl>
## 1 042312734      9.68            7.33           10.4
## 2 131882106      4.66            3.91            6.48
## 3 132584273      7.02            7.46            7.71
## 4 132642091      0              0              0
## 5 132685755     45.4            1.54          129.
## 6 132947386      4.08            4.07            8.90
## 7 210732575     14.3            0              22.9
## 8 231629970     -0.685          0.0719          5.43
## 9 237101094      3.84            5.84           12.0
## 10 237247909     11.8            6.39           15.0
## # i 36 more rows
```

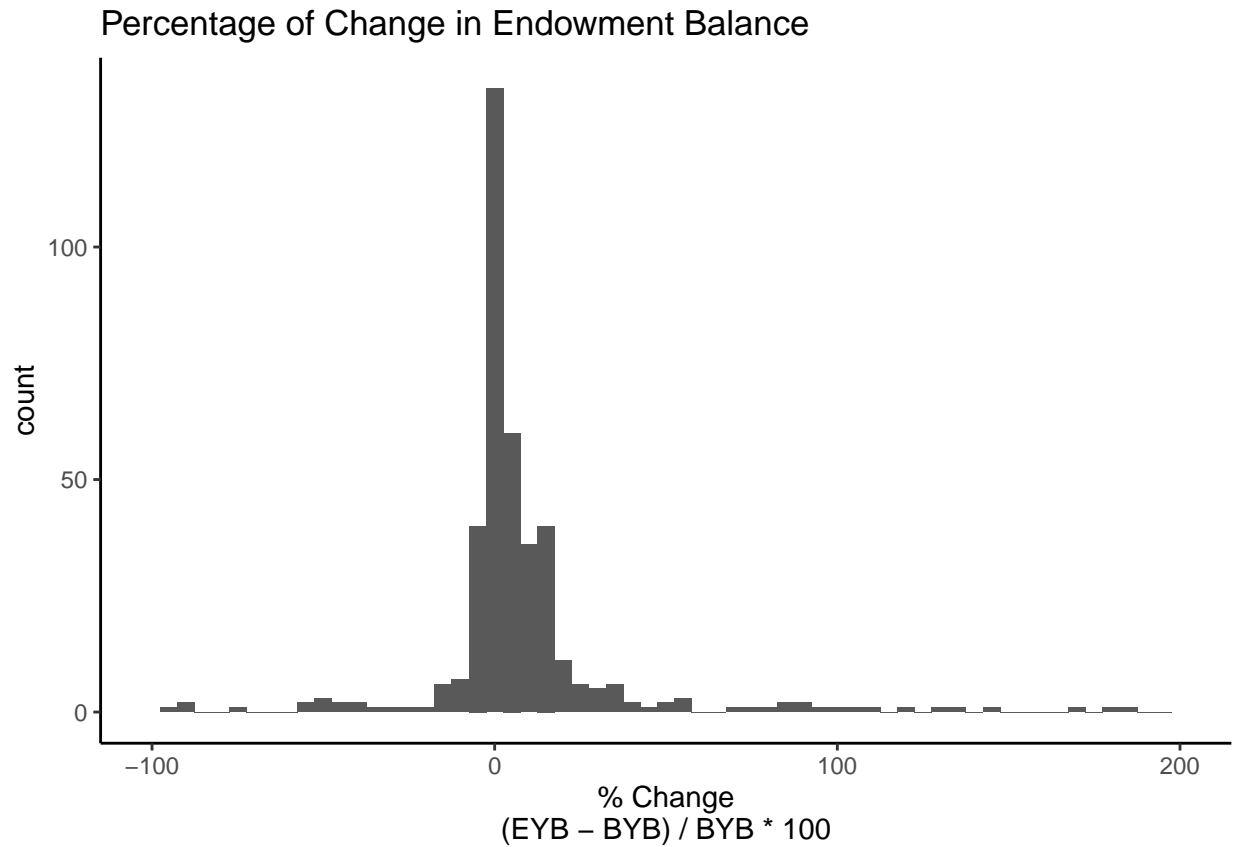
```
# Basic histogram summarizing it
ggplot(pct_change_ds, aes(x = pct_change)) +
  geom_histogram(binwidth = 20) +
  xlab("% Change\n(EYB - BYB) / BYB * 100") +
  ggtitle(label = "Percentage of Change in Endowment Balance", subtitle = "Red Line indicates 100%") +
  theme_classic() +
  geom_vline(xintercept = 100, color = "maroon", linetype = "dotted")
```

Percentage of Change in Endowment Balance

Red Line indicates 100%



```
## Histogram with cutoff of 200%
ggplot(pct_change_ds, aes(x = pct_change)) +
  geom_histogram(binwidth = 5) +
  xlab("% Change\n(EYB - BYB) / BYB * 100") +
  xlim(-100, 200) +
  ggtitle(label = "Percentage of Change in Endowment Balance") +
  theme_classic()
```



```
pct_change_ds %>%
  filter(pct_change != Inf) %>%
  select(organization_name, fiscal_year, pct_change) %>%
  make_table(title = "Percentage of Change in Endowment Balance", col_names = c("Name", "Fiscal Year", "Pct Change"),
    scroll_box(height = "450px"))
```

```
\begin{table}
\caption{
Percentage of Change in Endowment Balance
```

```
}
```

| Name | Fiscal Year | % Change |
|----------------------------|-------------|--------------|
| Joffrey Ballet | 2015 | 3091.4016854 |
| Fort Wayne Ballet | 2014 | 1897.2429619 |
| First State Ballet Theatre | 2012 | 1448.0762683 |
| Ballet Arizona | 2016 | 586.1374894 |
| Orlando Ballet | 2017 | 552.8554949 |
| Ballet Hispánico | 2021 | 432.1555786 |
| Nashville Ballet | 2011 | 288.7391599 |
| First State Ballet Theatre | 2020 | 242.6848638 |
| Milwaukee Ballet | 2011 | 212.5546600 |
| Atlanta Ballet | 2017 | 207.7002614 |
| Orlando Ballet | 2018 | 206.9063832 |
| Nashville Ballet | 2016 | 206.2101382 |
| Grand Rapids Ballet | 2015 | 186.0427033 |
| Richmond Ballet | 2016 | 178.0161483 |
| Charlotte Ballet | 2013 | 171.7520288 |
| Joffrey Ballet | 2019 | 146.9046722 |
| Texas Ballet Theater | 2015 | 133.5240000 |
| The Charleston Ballet | 2013 | 132.1823138 |
| Dayton Ballet | 2018 | 118.1185754 |
| Ballet Austin | 2013 | 111.9610260 |
| Miami City Ballet | 2021 | 103.3796951 |
| Ballet Memphis | 2018 | 100.3575811 |
| Ballet Memphis | 2012 | 94.1089954 |
| Joffrey Ballet | 2020 | 90.8216917 |
| Ballet Austin | 2017 | 90.2783217 |
| Texas Ballet Theater | 2018 | 85.8401625 |
| BalletMet | 2018 | 85.0249761 |
| Atlanta Ballet | 2016 | 81.8387443 |
| The Tallahassee Ballet | 2011 | 74.5093458 |
| First State Ballet Theatre | 2019 | 69.9900000 |
| Richmond Ballet | 2018 | 57.4030070 |
| Ballet Hispánico | 2014 | 54.7692390 |
| First State Ballet Theatre | 2016 | 53.2556470 |
| American Repertory Ballet | 2014 | 51.6506759 |
| Atlanta Ballet | 2012 | 47.7865882 |
| Dayton Ballet | 2019 | 47.3628143 |
| Ballet Des Moines | 2018 | 40.4800000 |
| Richmond Ballet | 2020 | 39.8449675 |
| The Sarasota Ballet | 2016 | 36.6266667 |
| Boston Ballet | 2014 | 36.2295633 |
| Eugene Ballet | 2021 | 35.3676599 |
| The Tallahassee Ballet | 2014 | 35.2860183 |
| Richmond Ballet | 2019 | 34.8263002 |
| American Repertory Ballet | 2013 | 34.4003625 |
| First State Ballet Theatre | 2017 | 32.2270270 |
| Nashville Ballet | 2013 | 31.9611808 |
| Ballet Memphis | 2014 | 29.0814635 |
| New Mexico Ballet Company | 2019 | 28.7289611 |
| Aspen Santa Fe Ballet | 2017 | 27.9853981 |
| Richmond Ballet | 2017 | 27.4751409 |
| Miami City Ballet | 2018 | 26.9869980 |
| Ballet Memphis | 2017 | 26.3332143 |
| BalletMet | 2014 | 25.3577272 |
| Kansas City Ballet | 2011 | 23.4046965 |
| BalletMet | 2019 | 22.9766213 |
| Ballet Memphis | 2011 | 21.9748193 |
| Joffrey Ballet | 2016 | 20.8230683 |

\end{table}

```
pct_change_ds %>%
  filter(pct_change != Inf) %>%
  select(organization_name, fiscal_year, pct_change) %>%
  group_by(fiscal_year) %>%
  summarize(total = n(),
             avg = mean(pct_change),
             med = median(pct_change),
             sd = sd(pct_change)) %>%
  make_table(title = "Percentage of Change in Endowment Balance By Year", col_names = c("Fiscal Year",
  scroll_box(height = "450px")
```

\begin{table}

\caption{

Percentage of Change in Endowment Balance By Year

}

| Fiscal Year | Total Companies | Average % Change | Median % Change | Standard Deviation |
|-------------|-----------------|------------------|-----------------|--------------------|
| 2010 | 1 | 0.3281092 | 0.3281092 | NA |
| 2011 | 31 | 23.2042543 | 8.9183455 | 64.52232 |
| 2012 | 36 | 39.5711928 | -1.2328310 | 242.88805 |
| 2013 | 40 | 14.4544304 | 4.8542974 | 38.48890 |
| 2014 | 40 | 56.4432194 | 8.2586243 | 299.10443 |
| 2015 | 41 | 83.4793525 | 0.0000000 | 482.88546 |
| 2016 | 43 | 24.2872831 | -0.7510343 | 98.50785 |
| 2017 | 43 | 27.5778232 | 8.6823347 | 89.08538 |
| 2018 | 44 | 15.9080630 | 4.8267305 | 46.13128 |
| 2019 | 43 | 5.6928617 | 1.1650569 | 32.37300 |
| 2020 | 39 | 7.2725415 | 0.2184921 | 44.93623 |
| 2021 | 6 | 98.6421536 | 27.8120030 | 167.76683 |

\end{table}

REMOVING investments

```
# Basic summary stats
pct_change_inv_ds %>%
  filter(!is.na(pct_change) & pct_change != Inf) %>%
  summarize(avg_pct_change = mean(pct_change),
             median_pct_change = median(pct_change),
             sd_pct_change = sd(pct_change))
```

```
## # A tibble: 1 x 3
##   avg_pct_change median_pct_change sd_pct_change
##   <dbl>             <dbl>             <dbl>
## 1      25.7          -0.685             199.
```

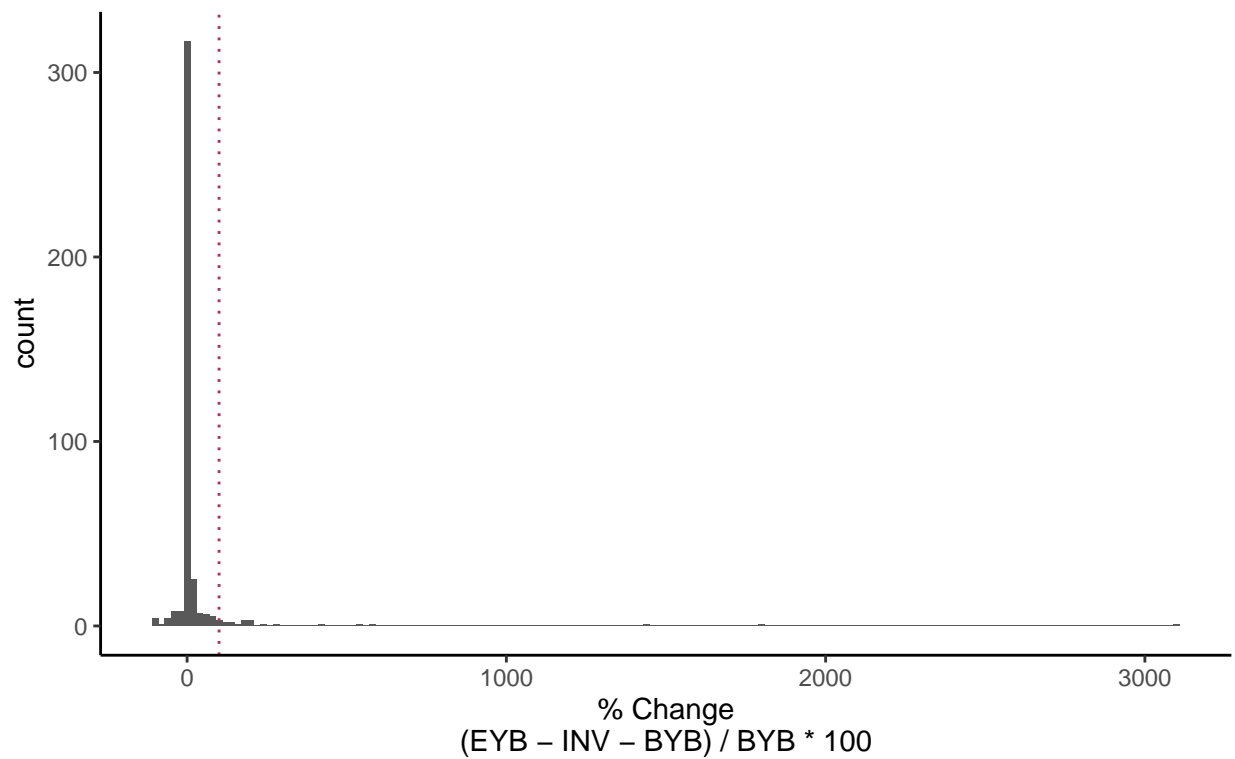
```
pct_change_inv_ds %>%
  filter(!is.na(pct_change) & pct_change != Inf) %>%
  group_by(EIN) %>%
  summarize(avg_pct_change = mean(pct_change),
            median_pct_change = median(pct_change),
            sd_pct_change = sd(pct_change))
```

```
## # A tibble: 46 x 4
##   EIN      avg_pct_change median_pct_change sd_pct_change
##   <chr>          <dbl>          <dbl>          <dbl>
## 1 042312734      1.83          -0.0413         7.19
## 2 131882106     -3.40          -3.70          4.03
## 3 132584273      1.27          -0.347         5.45
## 4 132642091     -0.0116         0            0.0187
## 5 132685755     39.3          -3.54        129.
## 6 132947386     -2.10          -3.28          3.68
## 7 210732575     14.3           0            22.9
## 8 231629970     -6.02          -3.99          6.15
## 9 237101094     -4.36          -6.15          5.04
## 10 237247909      6.51          -0.708        12.8
## # i 36 more rows
```

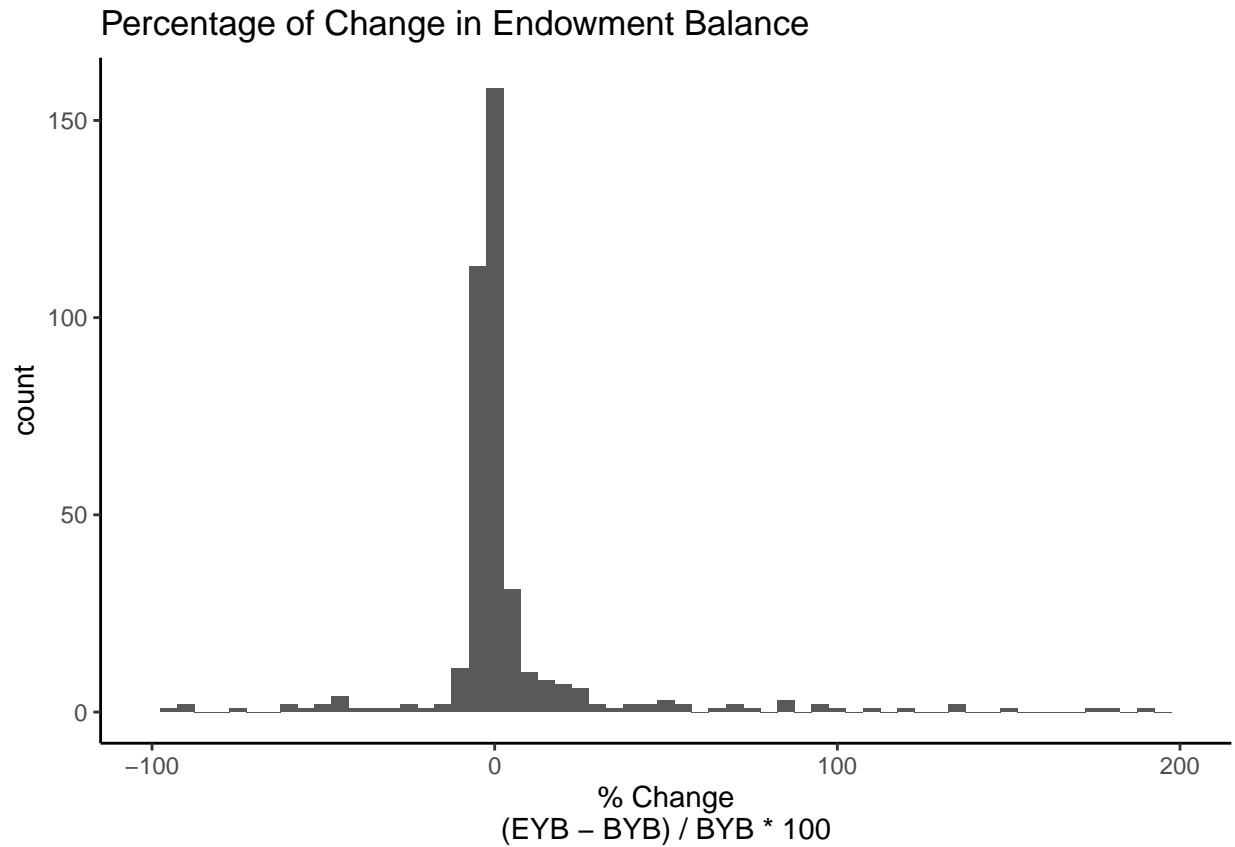
```
# Basic histogram summarizing it
ggplot(pct_change_inv_ds, aes(x = pct_change)) +
  geom_histogram(binwidth = 20) +
  xlab("% Change\n(EYB - INV - BYB) / BYB * 100") +
  ggtitle(label = "Percentage of Change in Endowment Balance", subtitle = "Red Line indicates 100%") +
  theme_classic() +
  geom_vline(xintercept = 100, color = "maroon", linetype = "dotted")
```


Percentage of Change in Endowment Balance

Red Line indicates 100%



```
## Histogram with cutoff of 200%
ggplot(pct_change_inv_ds, aes(x = pct_change)) +
  geom_histogram(binwidth = 5) +
  xlab("% Change\n(EYB - BYB) / BYB * 100") +
  xlim(-100, 200) +
  ggtitle(label = "Percentage of Change in Endowment Balance") +
  theme_classic()
```



```
pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  select(organization_name, fiscal_year, pct_change) %>%
  make_table(title = "Percentage of Change in Endowment Balance", col_names = c("Name", "Fiscal Year", "Pct Change"),
  scroll_box(height = "450px"))
```

```
\begin{table}
\caption{
Percentage of Change in Endowment Balance
}
```

| Name | Fiscal Year | % Change |
|----------------------------|-------------|--------------|
| Joffrey Ballet | 2015 | 3091.4016854 |
| Fort Wayne Ballet | 2014 | 1803.3224138 |
| First State Ballet Theatre | 2012 | 1448.0762683 |
| Ballet Arizona | 2016 | 586.1374894 |
| Orlando Ballet | 2017 | 547.7003508 |
| Ballet Hispánico | 2021 | 426.7571012 |
| Nashville Ballet | 2011 | 275.0123756 |
| First State Ballet Theatre | 2020 | 242.6848638 |
| Orlando Ballet | 2018 | 206.4099099 |
| Atlanta Ballet | 2017 | 202.3164983 |
| Milwaukee Ballet | 2011 | 197.6406871 |
| Nashville Ballet | 2016 | 187.7731485 |
| Grand Rapids Ballet | 2015 | 181.5004969 |
| Richmond Ballet | 2016 | 175.9261036 |
| Charlotte Ballet | 2013 | 151.2357429 |
| Joffrey Ballet | 2019 | 133.5419677 |
| Texas Ballet Theater | 2015 | 133.3333333 |
| The Charleston Ballet | 2013 | 118.3461667 |
| Dayton Ballet | 2018 | 111.2870455 |
| Ballet Memphis | 2012 | 99.9715525 |
| Ballet Austin | 2013 | 96.8010136 |
| Ballet Memphis | 2018 | 95.9769923 |
| Joffrey Ballet | 2020 | 86.1363009 |
| Texas Ballet Theater | 2018 | 85.6027576 |
| BalletMet | 2018 | 85.0249761 |
| Atlanta Ballet | 2016 | 76.1344084 |
| Ballet Austin | 2017 | 72.1291716 |
| First State Ballet Theatre | 2019 | 69.9900000 |
| Miami City Ballet | 2021 | 65.4636289 |
| The Tallahassee Ballet | 2011 | 56.0747664 |
| First State Ballet Theatre | 2016 | 53.2556470 |
| American Repertory Ballet | 2014 | 51.6506759 |
| Richmond Ballet | 2018 | 51.2753402 |
| Atlanta Ballet | 2012 | 47.9686134 |
| Dayton Ballet | 2019 | 45.1776417 |
| Richmond Ballet | 2020 | 43.9335422 |
| Ballet Hispánico | 2014 | 41.4922805 |
| The Sarasota Ballet | 2016 | 37.6855000 |
| American Repertory Ballet | 2013 | 34.4003625 |
| First State Ballet Theatre | 2017 | 32.2270270 |
| Aspen Santa Fe Ballet | 2017 | 27.9853981 |
| Richmond Ballet | 2019 | 26.8630003 |
| BalletMet | 2014 | 25.3577272 |
| Ballet Des Moines | 2018 | 25.2500000 |
| Joffrey Ballet | 2016 | 24.4935699 |
| BalletMet | 2019 | 22.9766213 |
| Richmond Ballet | 2017 | 22.7958845 |
| Kansas City Ballet | 2012 | 21.4853681 |
| New Mexico Ballet Company | 2019 | 21.2777356 |
| Miami City Ballet | 2018 | 20.0147230 |
| Ballet Memphis | 2014 | 19.3704481 |
| Boston Ballet | 2014 | 18.6852707 |
| Nashville Ballet | 2013 | 18.4613232 |
| The Tallahassee Ballet | 2014 | 18.0748529 |
| Ballet Memphis | 2017 | 16.5882353 |
| The Tallahassee Ballet | 2015 | 15.7290896 |
| Oregon Ballet Theatre | 2013 | 15.5103005 |

\end{table}

```
pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  select(organization_name, fiscal_year, pct_change) %>%
  group_by(fiscal_year) %>%
  summarize(total = n(),
            avg = mean(pct_change),
            med = median(pct_change),
            sd = sd(pct_change)) %>%
  make_table(title = "Percentage of Change in Endowment Balance By Year", col_names = c("Fiscal Year",
  scroll_box(height = "450px")
```

\begin{table}

\caption{

Percentage of Change in Endowment Balance By Year

}

| Fiscal Year | Total Companies | Average % Change | Median % Change | Standard Deviation |
|-------------|-----------------|------------------|-----------------|--------------------|
| 2010 | 1 | -7.337498 | -7.3374978 | NA |
| 2011 | 31 | 13.177900 | -1.5332290 | 62.57769 |
| 2012 | 36 | 39.939578 | -1.1673363 | 242.90538 |
| 2013 | 40 | 7.060537 | -1.6824371 | 35.57750 |
| 2014 | 40 | 45.612283 | -0.1755000 | 285.55196 |
| 2015 | 41 | 81.554256 | -0.2569099 | 483.16993 |
| 2016 | 43 | 24.154434 | -0.4368254 | 97.43600 |
| 2017 | 43 | 19.912496 | 0.0000000 | 89.40665 |
| 2018 | 44 | 11.162350 | -0.4508906 | 46.15232 |
| 2019 | 43 | 1.888381 | -2.4754845 | 31.06880 |
| 2020 | 39 | 5.044147 | -1.0929032 | 45.02615 |
| 2021 | 6 | 81.164577 | 0.2636754 | 171.41046 |

\end{table}

Range of Endowment Percent Change

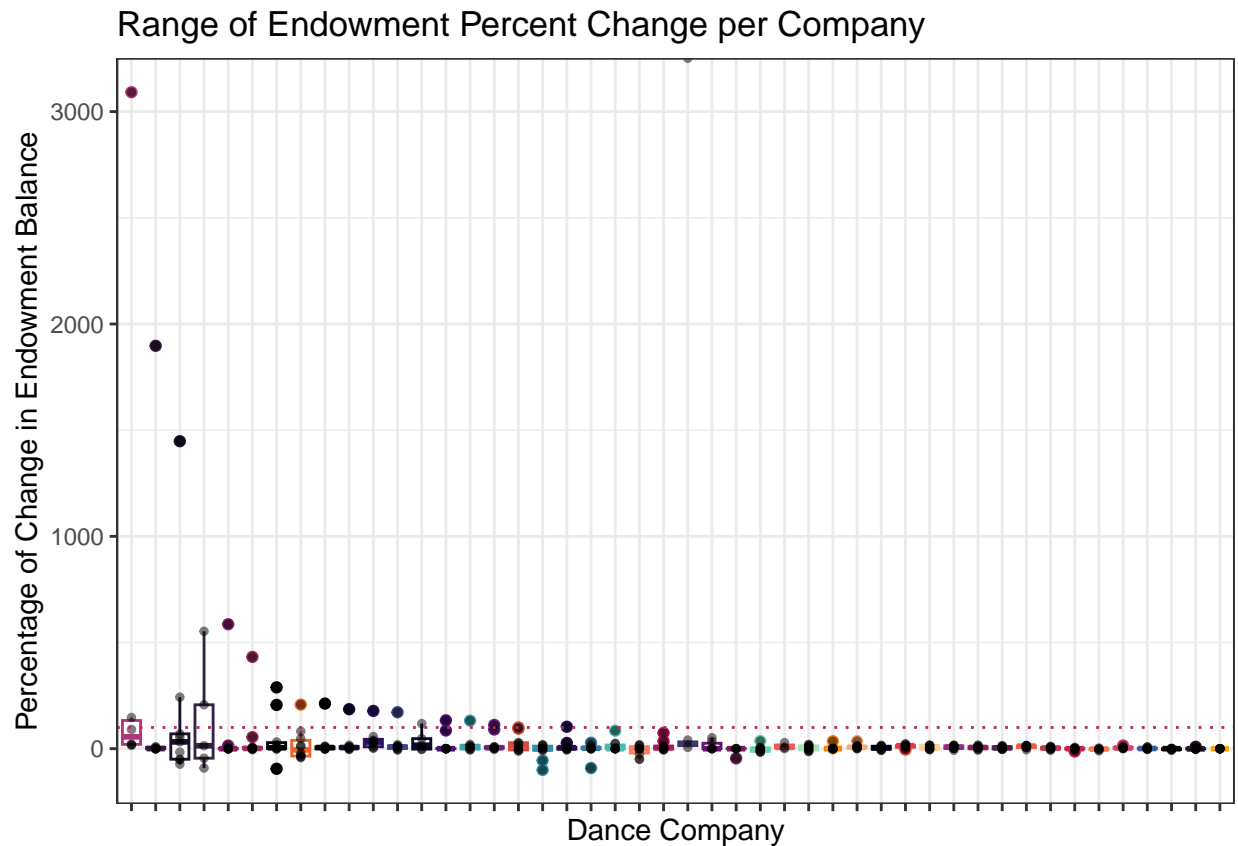
WITH investments

```
## Ranges of different percent change
# Reordering by standard deviation of pct_spend down
pct_change_ds_box <- pct_change_ds %>%
  group_by(organization_name) %>%
  filter(pct_change != Inf) %>%
  summarize(sd = sd(pct_change, na.rm = TRUE)) %>%
  right_join(pct_change_ds, by = "organization_name") %>%
  select(organization_name, EIN, pct_change, sd) %>%
  mutate(organization_name = reorder(organization_name, -sd, na.rm = TRUE))
## Unlimited
box_plot <- ggplot(pct_change_ds_box, aes(x = organization_name, y = pct_change)) +
```

```

geom_boxplot(aes(color = organization_name), show.legend = FALSE) +
geom_point(size = 1, alpha = 0.5, show.legend = FALSE) +
theme_bw() +
labs(title = "Range of Endowment Percent Change per Company",
     x = "Dance Company",
     y = "Percentage of Change in Endowment Balance") +
theme(axis.text.x = element_blank()) +
geom_hline(yintercept = 100, linetype = "dotted", color = "maroon") +
scale_colour_manual(values = pal)
box_plot

```

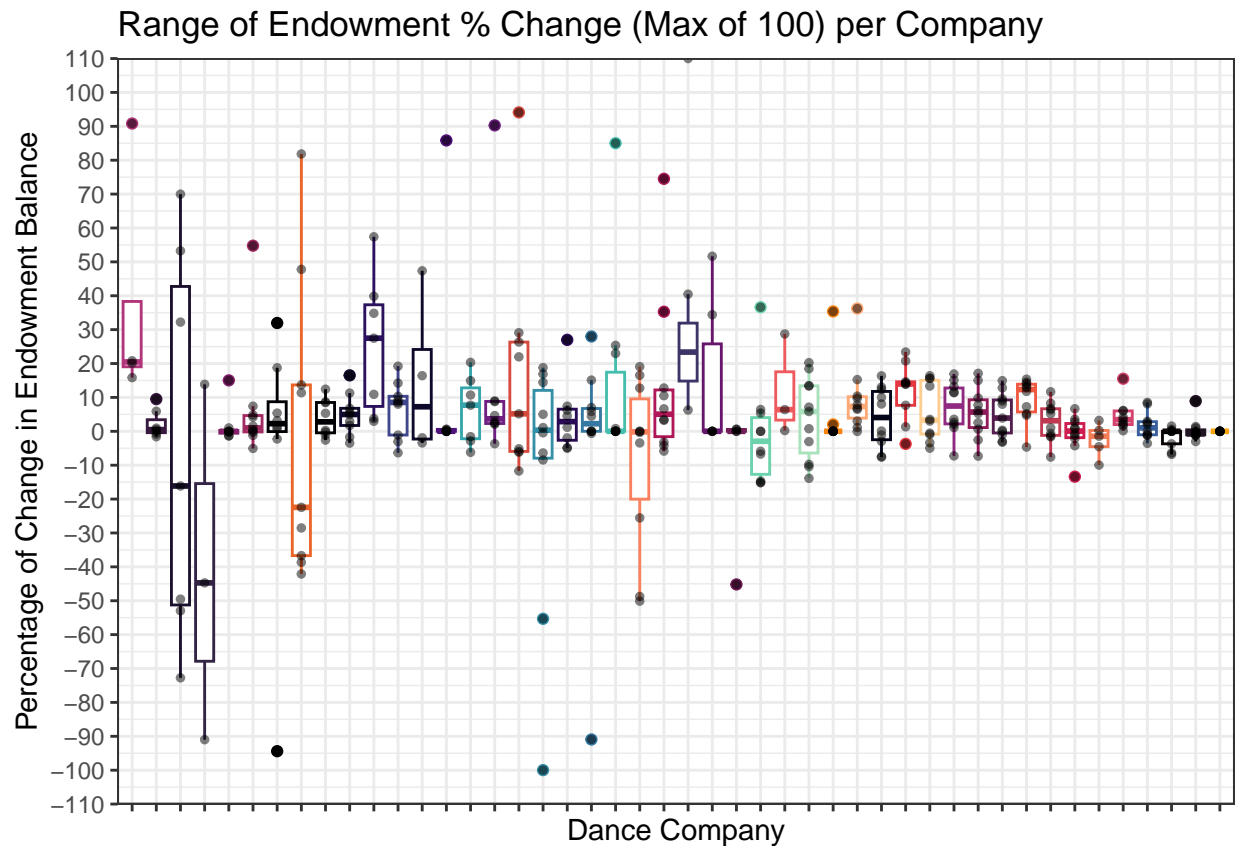


```

#ggplotly(box_plot) %>%
# layout(showlegend = FALSE)
##Limited to 100 for visibility
box_plot_lim <- ggplot(pct_change_ds_box, aes(x = organization_name, y = pct_change)) +
  geom_boxplot(aes(color = organization_name), show.legend = FALSE) +
  geom_point(size = 1, alpha = 0.5) +
  theme_bw() +
  labs(title = "Range of Endowment % Change (Max of 100) per Company",
       x = "Dance Company",
       y = "Percentage of Change in Endowment Balance") +
  theme(axis.text.x = element_blank()) +
  scale_y_continuous(breaks = scales::breaks_pretty(n = 20),
                    limit = c(-100, 100)) +
  scale_colour_manual(values = pal)

```

```
box_plot_lim
```



```
#ggplotly(box_plot_lim) %>%
#layout(showlegend = FALSE)
```

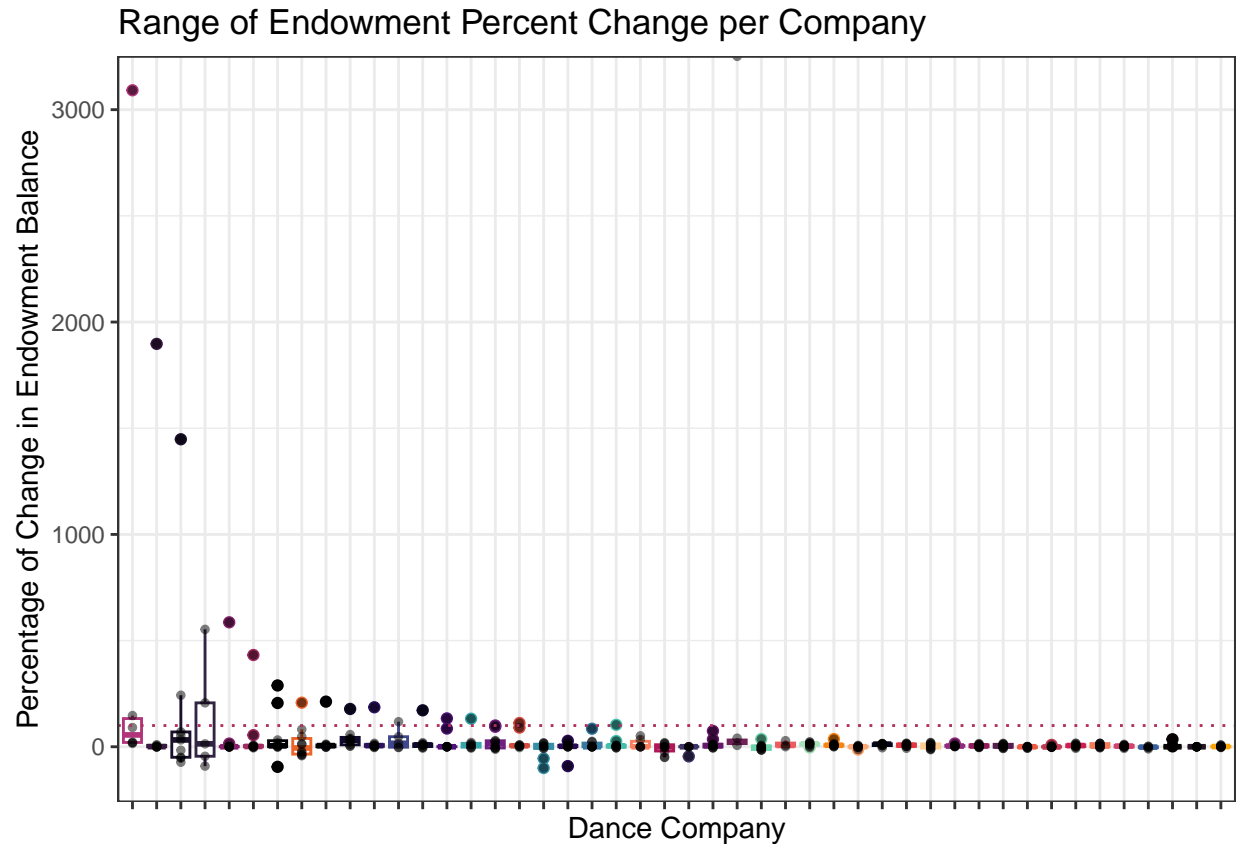
WITHOUT investments

```
## Ranges of different percent change
# reorder(organization_name, pull(summarize(spend_down, sd = sd(group_by(spend_down, pct_change))))), na
# Reordering by standard deviation of pct_spend down
pct_change_inv_box <- pct_change_inv_ds %>%
  group_by(organization_name) %>%
  filter(pct_change != Inf) %>%
  summarize(sd = sd(pct_change, na.rm = TRUE)) %>%
  right_join(pct_change_ds, by = "organization_name") %>%
  select(organization_name, EIN, pct_change, sd) %>%
  mutate(organization_name = reorder(organization_name, -sd, na.rm = TRUE))
## Unlimited
box_plot_inv <- ggplot(pct_change_inv_box, aes(x = organization_name, y = pct_change)) +
  geom_boxplot(aes(color = organization_name), show.legend = FALSE) +
  geom_point(size = 1, alpha = 0.5) +
  theme_bw() +
  labs(title = "Range of Endowment Percent Change per Company",
```

```

x = "Dance Company",
y = "Percentage of Change in Endowment Balance") +
theme(axis.text.x = element_blank()) +
geom_hline(yintercept = 100, linetype = "dotted", color = "maroon") +
scale_colour_manual(values = pal)
box_plot_inv

```



```

#ggplotly(box_plot_inv) %>%
  #layout(showlegend = FALSE)
##Limited to 100 for visibility
box_plot_inv_lim <- ggplot(pct_change_inv_box, aes(x = organization_name, y = pct_change)) +
  geom_boxplot(aes(color = organization_name), show.legend = FALSE) +
  geom_point(size = 1, alpha = 0.5) +
  theme_bw() +
  labs(title = "Range of Endowment % Change (Max of 100) per Company",
        x = "Dance Company",
        y = "Percentage of Change in Endowment Balance") +
  theme(axis.text.x = element_blank()) +
  scale_y_continuous(breaks = scales::breaks_pretty(n = 20),
                    limit = c(-100, 100)) +
  scale_colour_manual(values = pal)
#ggplotly(box_plot_inv_lim) %>%
  # layout(showlegend = FALSE)

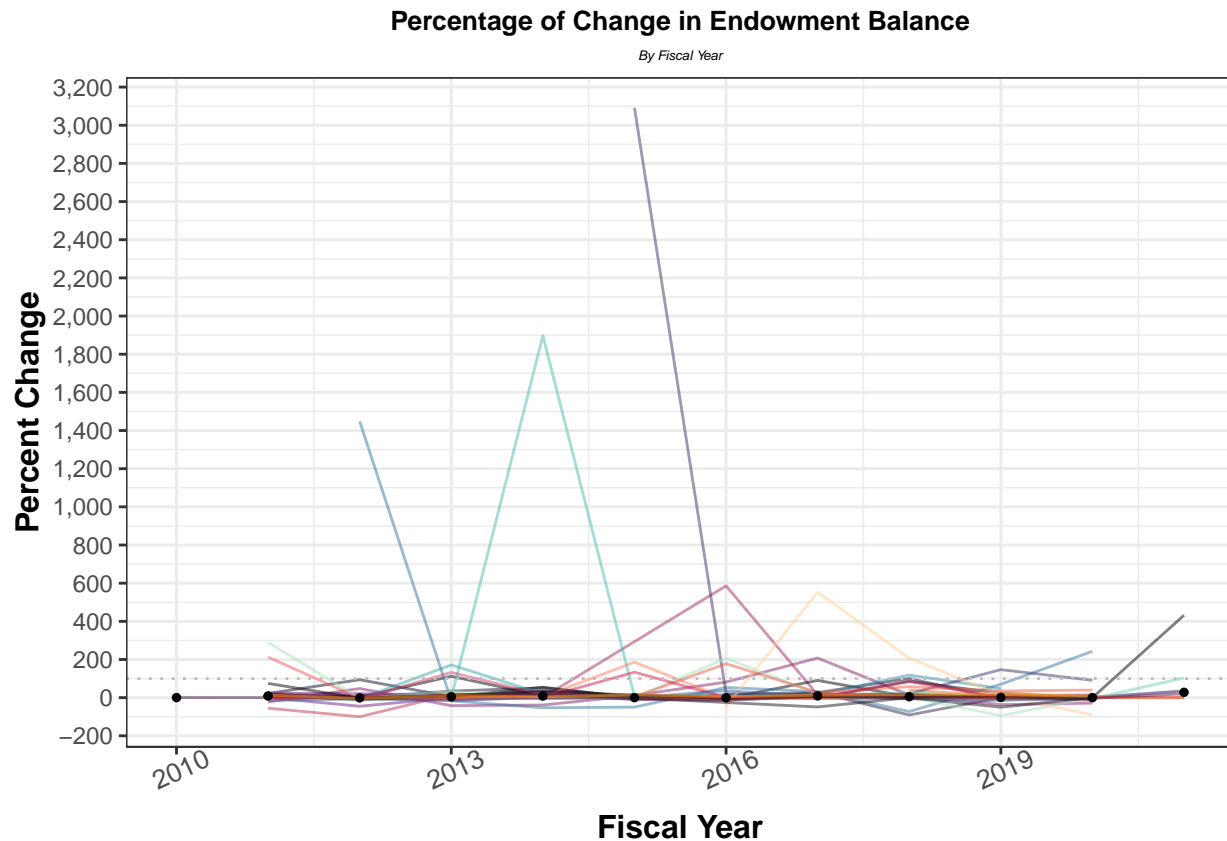
```

Percent Change over Time

WITH investments

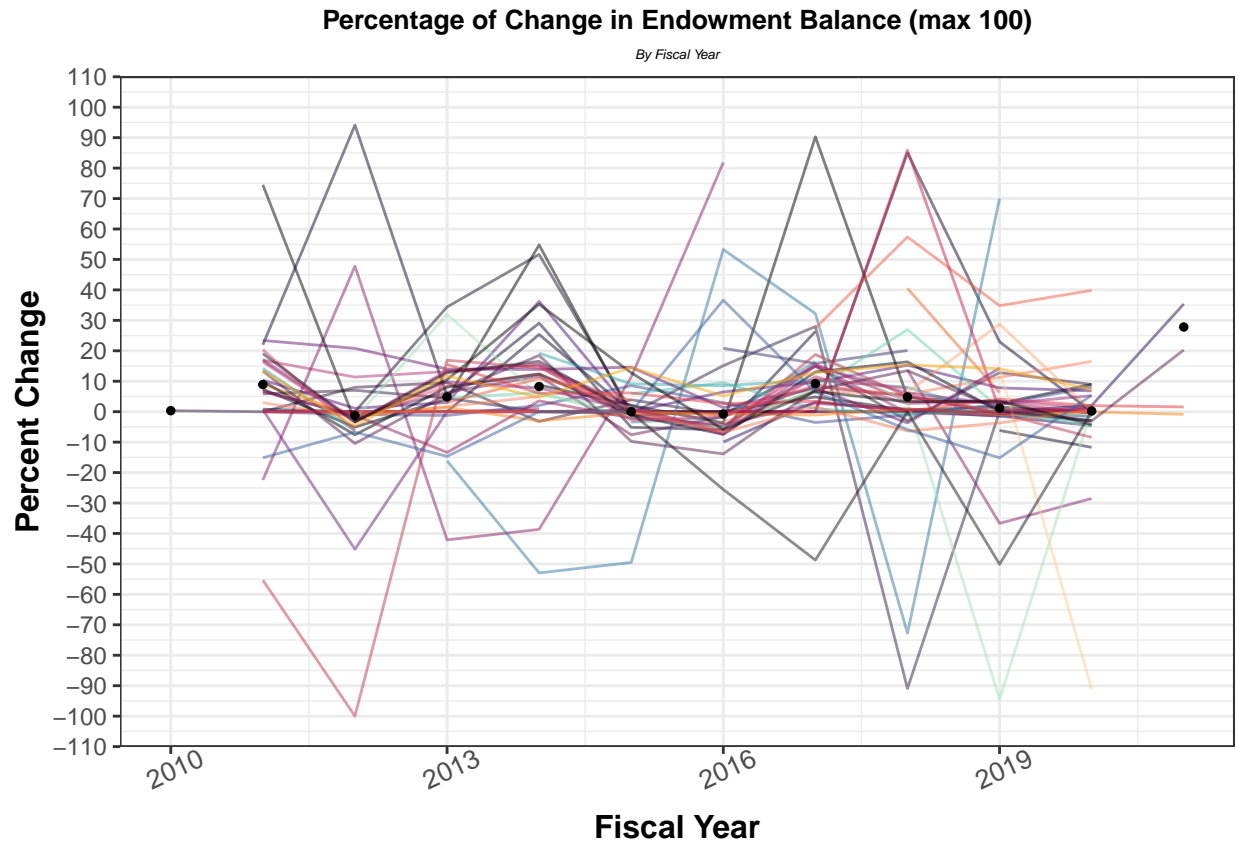
```
## Retrieving median values
pct_change_med <- pct_change_ds %>%
  group_by(fiscal_year) %>%
  summarize(median = median(pct_change, na.rm = TRUE)) %>%
  mutate(organization_name = "Median",
         fiscal_year_dt = paste(fiscal_year, "-01-01", sep = ""))
pct_change_med$fiscal_year_dt = as.Date(pct_change_med$fiscal_year_dt, format = "%Y-%m-%d")

## Spend Down over Time
pct_change_plot <- pct_change_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
            group = organization_name, color = organization_name)) +
  geom_line(alpha = 0.5, show.legend = FALSE) +
  theme_bw() +
  scale_colour_manual(values = pal) +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance",
       subtitle = "By Fiscal Year") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face = "bold", size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size = 7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 20)) +
  geom_hline(yintercept = 100, linetype = "dotted", color = "gray") +
  geom_point(data = pct_change_med, aes(x = fiscal_year, y = median), color = "black", size = 1)
pct_change_plot
```

```
#ggplotly(pct_change_plot)
```

```
##Plot with Y scale between -100 and 100
limited_scale <- pct_change_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
             group = organization_name, color = organization_name)) +
  geom_line(show.legend = FALSE, alpha = 0.5) +
  theme_bw() +
  scale_colour_manual(values = pal) +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance (max 100)",
       subtitle = "By Fiscal Year") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 20),
                    limits = c(-100, 100)) +
  geom_point(data = pct_change_med, aes(x = fiscal_year, y = median), color = "black", size = 1)
limited_scale
```



```
#ggplotly(limited_scale)
```

WITHOUT investments

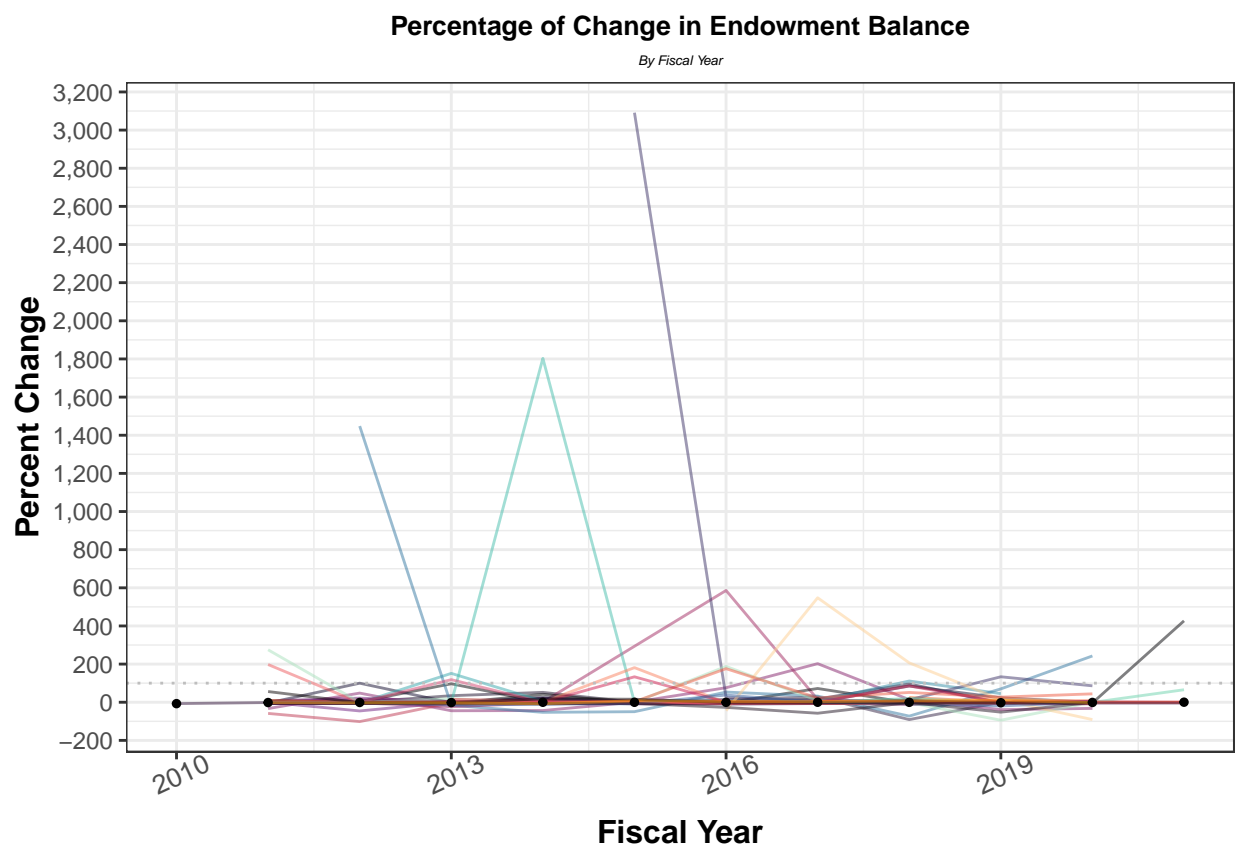
```
## Retrieving median values
pct_change_inv_med <- pct_change_inv_ds %>%
  group_by(fiscal_year) %>%
  summarize(median = median(pct_change, na.rm = TRUE)) %>%
  mutate(organization_name = "Median")

## Spend Down over Time
pct_change_inv_plot <- pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
             group = organization_name, color = organization_name)) +
  geom_line(alpha = 0.5, show.legend = FALSE) +
  scale_colour_manual(values = pal) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance",
```

```

    subtitle = "By Fiscal Year") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 20)) +
  geom_hline(yintercept = 100, linetype = "dotted", color = "gray") +
  geom_point(data = pct_change_inv_med, aes(x = fiscal_year, y = median), color = "black", size = 1)
pct_change_inv_plot

```



```
#ggplotly(pct_change_inv_plot)
```

```

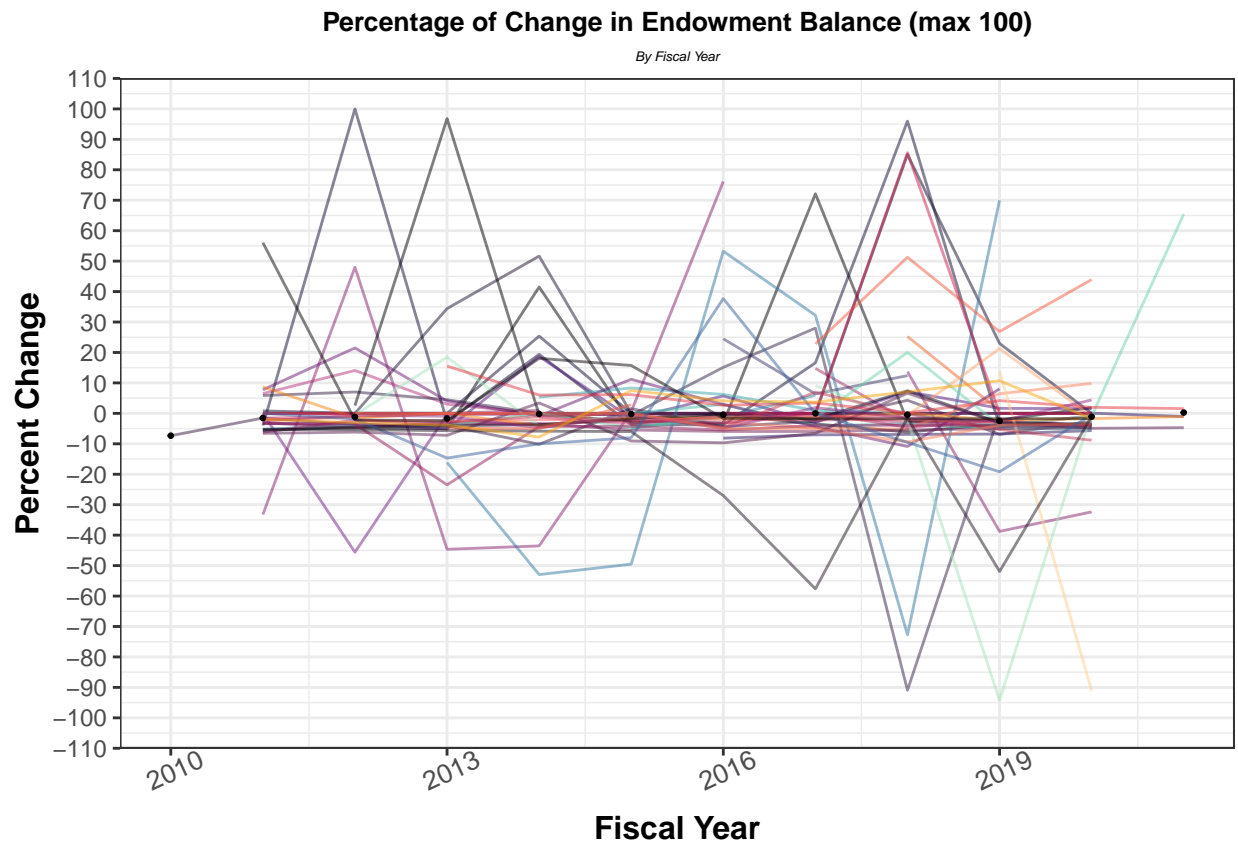
##Plot with Y scale between -100 and 100
limited_scale_inv <- pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
            group = organization_name, color = organization_name)) +
  geom_line(show.legend = FALSE, alpha = 0.5) +
  scale_colour_manual(values = pal) +
  theme_bw() +
  labs(y = "Percent Change",

```

```

x = "Fiscal Year",
title = "Percentage of Change in Endowment Balance (max 100)",
subtitle = "By Fiscal Year") +
theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
axis.title = element_text(size = 12, face = "bold"),
plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
axis.text.x = element_text(size = 10, angle = 25),
strip.text = element_text(face="bold",size = 5),
legend.key.size = unit(1, 'mm'),
legend.text = element_text(size=7)) +
scale_y_continuous(labels = scales::comma_format(),
breaks = scales::pretty_breaks(n = 20),
limits = c(-100, 100)) +
geom_point(data = pct_change_inv_med, aes(x = fiscal_year, y = median), color = "black", size = 0.5)
limited_scale_inv

```



```
#ggplotly(limited_scale_inv)
```

Comparing within-year percent change to S&P 500

```
## Adding end date
source(here("GET_VARS.R"))
```

```

##Specifically reading in Tax Period End Date
files <- dir( here("ballet_990_released_20230208"),
             full.names = TRUE)
taxperiod_data <- map_df(files, ~
  get_df(variables = c("//Return//ReturnHeader//TaxPeriodEndDt"),
          filename = .x
        ))
taxperiod_data$fiscal_year = as.numeric(as.character(taxperiod_data$fiscal_year))
pct_change_ds <- pct_change_ds %>%
  left_join(taxperiod_data, by = c("EIN", "fiscal_year"))
pct_change_inv_ds <- pct_change_inv_ds %>%
  left_join(taxperiod_data, by = c("EIN", "fiscal_year"))
## Imputing return_date from N/A
# Assuming that month/day is the same, fiscal year is one year prior
pct_change_ds <- pct_change_ds %>%
  group_by(EIN) %>%
  arrange(desc(EIN), desc(fiscal_year)) %>%
  mutate(TaxPeriodEndDt= ifelse(is.na(TaxPeriodEndDt),
                                paste(fiscal_year, substring(lag(TaxPeriodEndDt), 5, 10), sep = ""),
                                as.character(TaxPeriodEndDt)))
pct_change_inv_ds <- pct_change_inv_ds %>%
  group_by(EIN) %>%
  arrange(desc(EIN), desc(fiscal_year)) %>%
  mutate(TaxPeriodEndDt= ifelse(is.na(TaxPeriodEndDt),
                                paste(fiscal_year, substring(lag(TaxPeriodEndDt), 5, 10), sep = ""),
                                as.character(TaxPeriodEndDt)))
#Produces some NAs due to vectorization

# Rids all '2010NA' etc.
# Only have 11 years, so at max will run 11 times
for (i in 1:11) {
  pct_change_ds <- pct_change_ds %>%
    mutate(TaxPeriodEndDt= ifelse(grepl("NA", TaxPeriodEndDt),
                                  paste(fiscal_year, substring(lag(TaxPeriodEndDt), 5, 10), sep = ""),
                                  as.character(TaxPeriodEndDt)))

  pct_change_inv_ds <- pct_change_inv_ds %>%
    mutate(TaxPeriodEndDt= ifelse(grepl("NA", TaxPeriodEndDt),
                                  paste(fiscal_year, substring(lag(TaxPeriodEndDt), 5, 10), sep = ""),
                                  as.character(TaxPeriodEndDt)))
}
pct_change_ds$TaxPeriodEndDt = as.Date(pct_change_ds$TaxPeriodEndDt, format = "%Y-%m-%d")
pct_change_inv_ds$TaxPeriodEndDt = as.Date(pct_change_inv_ds$TaxPeriodEndDt, format = "%Y-%m-%d")
pct_change_ds <- ungroup(pct_change_ds)
pct_change_inv_ds <- ungroup(pct_change_inv_ds)

```

WITH investments

```

## GETTING PERCENT CHANGE OF FISCAL YEARS THAT EXIST IN THE DATASET
## Year wise PCT change for S&P 500
sp_end <- read_csv(here('data', 'SP500.csv')) %>%

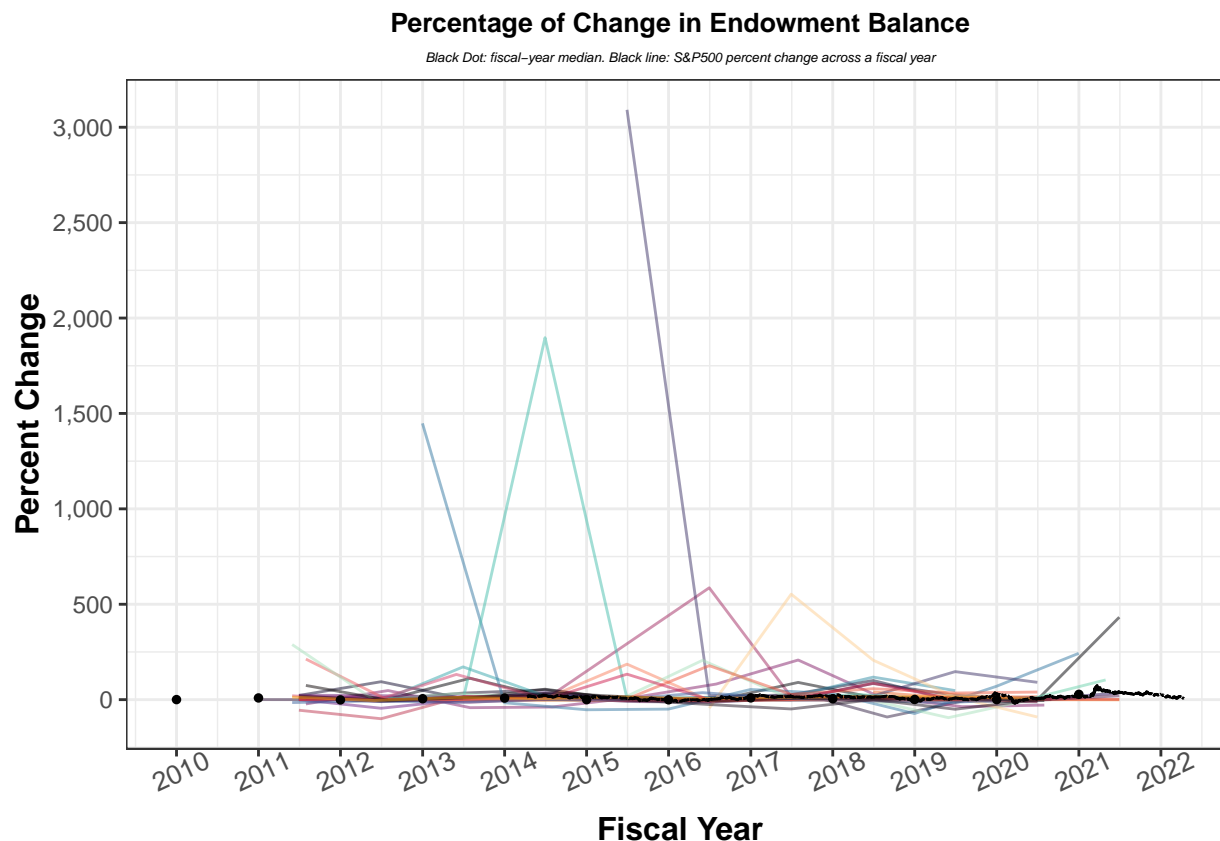
```

```

rename_with(tolower) %>%
mutate(sp500 = as.numeric(sp500)) %>%
filter(!is.na(sp500)) %>%
select(date_end = date, sp500_end = sp500) %>%
mutate(month = month(date_end),
       year = year(date_end),
       # Weekend / National holiday detection
       date_last = case_when(
         wday(date_end - 365) == as.factor("1") ~ date_end - 367, #Sunday
         wday(date_end - 365) == as.factor("7") ~ date_end - 366, #Saturday
         TRUE ~ date_end - 365
       ))
## All 'starting' fiscal year values
sp_start <- read_csv(here('data', 'SP500.csv')) %>%
  rename_with(tolower) %>%
  mutate(sp500 = as.numeric(sp500)) %>%
  filter(!is.na(sp500)) %>%
  select(date_start = date, sp500_start = sp500) %>%
  mutate(month = month(date_start),
         year = year(date_start))
#Change over the fiscal year
sp_change_years <- sp_end %>%
  left_join(sp_start, by = c("date_last" = "date_start")) %>%
  mutate(pct_change_yr = (sp500_end - sp500_start)/sp500_start * 100)

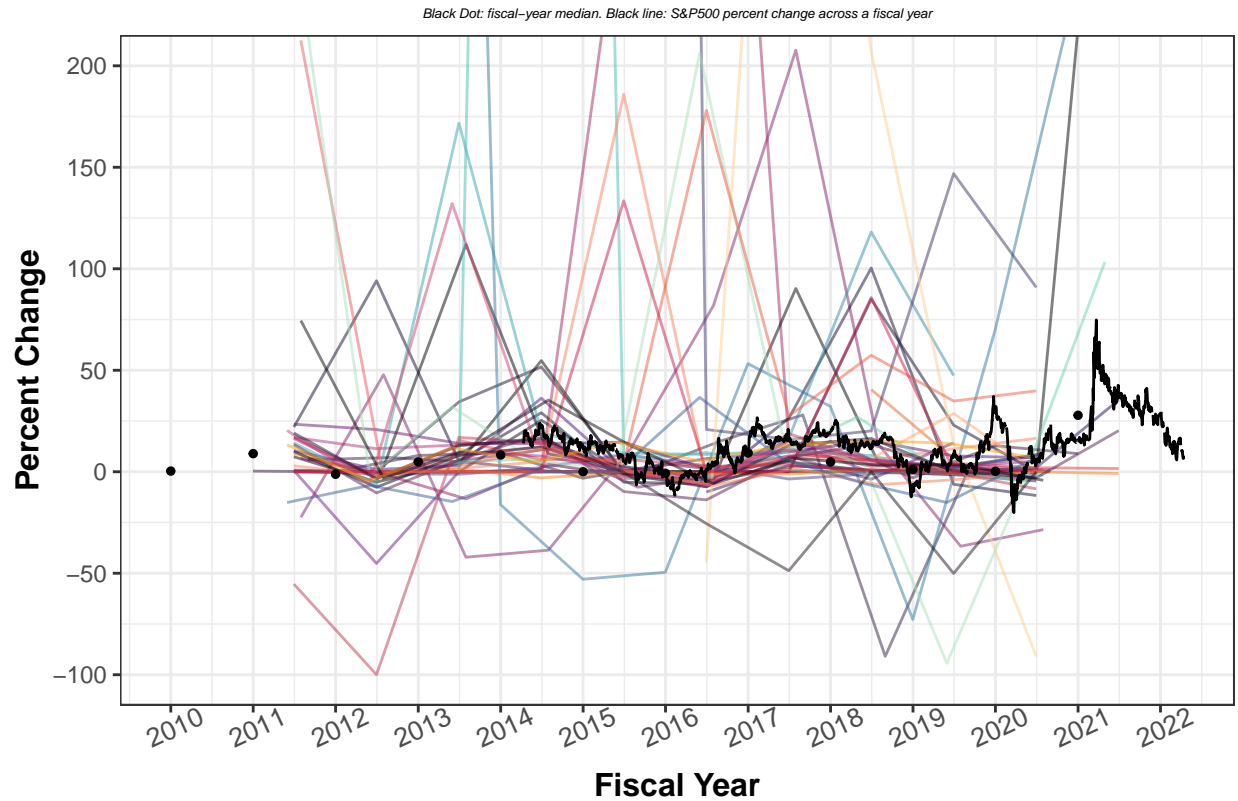
## Spend Down over Time
## With year-wise pct change for S&P 500
pc_endow_sp_fr <- pct_change_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = TaxPeriodEndDt, y = pct_change)) +
  geom_line(aes(group = organization_name, color = organization_name), alpha = 0.5, show.legend = FALSE) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance",
       subtitle = "Black Dot: fiscal-year median. Black line: S&P500 percent change across a fiscal year")
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  scale_color_manual(values=pal) +
  scale_x_date(breaks = scales::pretty_breaks(n = 20)) +
  geom_point(data = pct_change_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 1) +
  geom_line(data = sp_change_years, aes(x = date_end, y = pct_change_yr), color = "black", show.legend = FALSE)
pc_endow_sp_fr

```



```
##Limited to 200
pc_endow_sp_fr_200 <- pct_change_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = TaxPeriodEndDt, y = pct_change)) +
  geom_line(aes(group = organization_name, color = organization_name), alpha = 0.5, show.legend = FALSE) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance, Zoomed in to 200%",
       subtitle = "Black Dot: fiscal-year median. Black line: S&P500 percent change across a fiscal year")
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  scale_color_manual(values=pal) +
  coord_cartesian(ylim=c(-100, 200)) +
  scale_x_date(breaks = scales::pretty_breaks(n = 20)) +
  geom_point(data = pct_change_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 1) +
  geom_line(data = sp_change_years, aes(x = date_end, y = pct_change_yr), color = "black", show.legend = FALSE)
pc_endow_sp_fr_200
```

Percentage of Change in Endowment Balance, Zoomed in to 200%



```
pct_change_ds %>%
  filter(pct_change != Inf) %>%
  left_join(sp_change_years, by = c("TaxPeriodEndDt" = "date_end")) %>%
  select(organization_name, TaxPeriodEndDt, pct_change, pct_change_yr) %>%
  arrange(desc(abs(pct_change))) %>%
  make_table(title = "Percentage Change of Endowment and S&P 500",
             col_names = c("Organization Name", "Fiscal Year End Date", "Endowment Percent Change", "S&P 500 Percent Change"),
             scroll_box(height = "450px"))
```

\begin{table}

\caption{

Percentage Change of Endowment and S&P 500

}

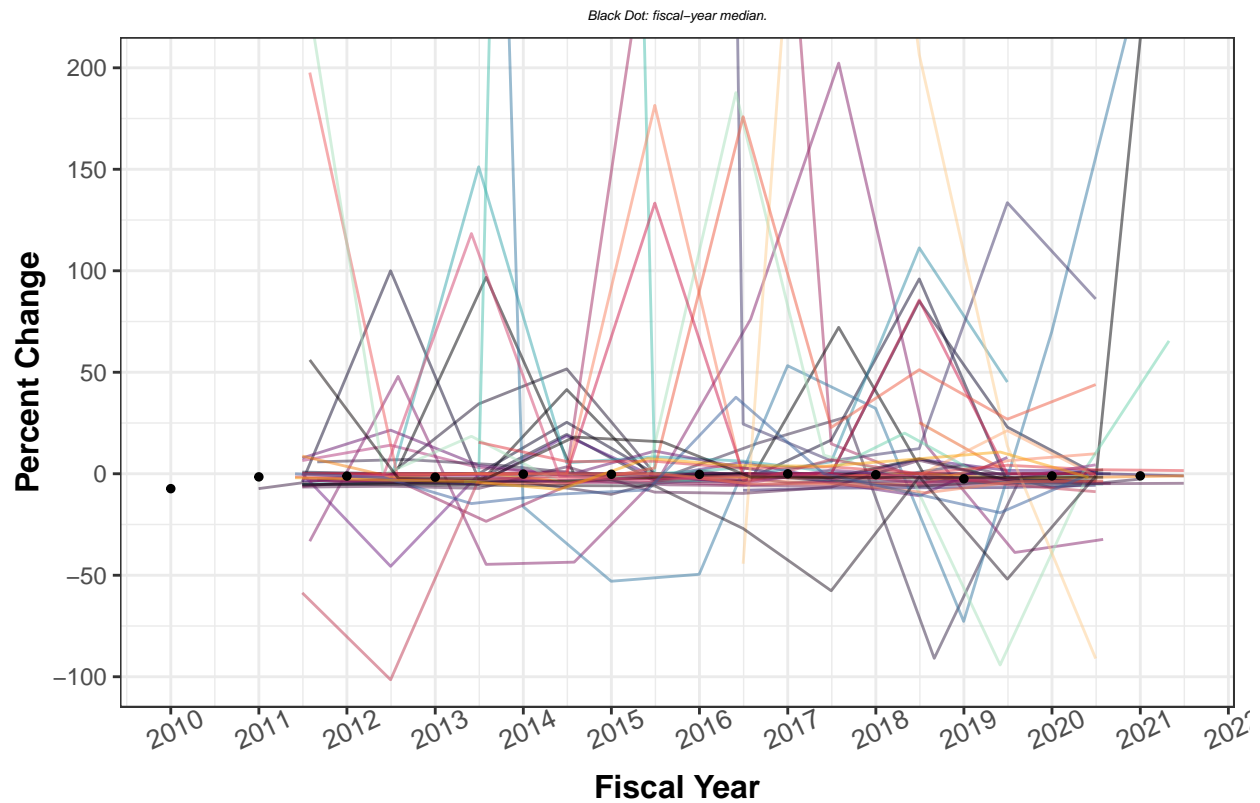
| Organization Name | Fiscal Year End Date | Endowment Percent Change | S&P 500 Percent |
|----------------------------|----------------------|--------------------------|-----------------|
| Joffrey Ballet | 2015-06-30 | 3091.4016854 | |
| Fort Wayne Ballet | 2014-06-30 | 1897.2429619 | 2 |
| First State Ballet Theatre | 2012-12-31 | 1448.0762683 | |
| Ballet Arizona | 2016-06-30 | 586.1374894 | |
| Orlando Ballet | 2017-06-30 | 552.8554949 | 1 |
| Ballet Hispánico | 2021-06-30 | 432.1555786 | 3 |
| Nashville Ballet | 2011-05-31 | 288.7391599 | |
| First State Ballet Theatre | 2020-12-31 | 242.6848638 | |
| First State Ballet Theatre | 2020-12-31 | 242.6848638 | |
| Milwaukee Ballet | 2011-07-31 | 212.5546600 | |
| Atlanta Ballet | 2017-07-31 | 207.7002614 | 1 |
| Orlando Ballet | 2018-06-30 | 206.9063832 | |
| Nashville Ballet | 2016-05-31 | 206.2101382 | - |
| Grand Rapids Ballet | 2015-06-30 | 186.0427033 | |
| Richmond Ballet | 2016-06-30 | 178.0161483 | |
| Charlotte Ballet | 2013-06-30 | 171.7520288 | |
| Joffrey Ballet | 2019-06-30 | 146.9046722 | |
| Texas Ballet Theater | 2015-06-30 | 133.5240000 | |
| The Charleston Ballet | 2013-05-31 | 132.1823138 | |
| Dayton Ballet | 2018-06-30 | 118.1185754 | |
| Ballet Austin | 2013-07-31 | 111.9610260 | |
| Miami City Ballet | 2021-04-30 | 103.3796951 | 4 |
| Ballet Memphis | 2018-06-30 | 100.3575811 | |
| San Francisco Ballet | 2012-06-30 | -99.9832016 | |
| Nashville Ballet | 2019-05-31 | -94.4004512 | |
| Ballet Memphis | 2012-06-30 | 94.1089954 | |
| Orlando Ballet | 2020-06-30 | -90.9983829 | |
| Aspen Santa Fe Ballet | 2018-08-31 | -90.9315940 | 1 |
| Joffrey Ballet | 2020-06-30 | 90.8216917 | |
| Ballet Austin | 2017-07-31 | 90.2783217 | 1 |
| Texas Ballet Theater | 2018-06-30 | 85.8401625 | |
| BalletMet | 2018-06-30 | 85.0249761 | |
| Atlanta Ballet | 2016-07-31 | 81.8387443 | |
| The Tallahassee Ballet | 2011-07-31 | 74.5093458 | |
| First State Ballet Theatre | 2018-12-31 | -72.7468454 | - |
| First State Ballet Theatre | 2019-12-31 | 69.9900000 | 2 |
| Richmond Ballet | 2018-06-30 | 57.4030070 | |
| San Francisco Ballet | 2011-06-30 | -55.3267726 | |
| Ballet Hispánico | 2014-06-30 | 54.7692390 | 2 |
| First State Ballet Theatre | 2016-12-31 | 53.2556470 | |
| First State Ballet Theatre | 2014-12-31 | -52.9562948 | 1 |
| American Repertory Ballet | 2014-06-30 | 51.6506759 | 2 |
| The Washington Ballet | 2019-06-30 | -50.1144953 | |
| First State Ballet Theatre | 2015-12-31 | -49.5289330 | - |
| The Washington Ballet | 2017-06-30 | -48.7379222 | 1 |
| Atlanta Ballet | 2012-07-31 | 47.7865882 | |
| Dayton Ballet | 2019-06-30 | 47.3628143 | |
| Colorado Ballet | 2012-06-30 | -45.1865576 | |
| Orlando Ballet | 2016-06-30 | -44.7242762 | |
| Atlanta Ballet | 2013-07-31 | -42.0972020 | |
| Ballet Des Moines | 2018-06-30 | 40.4800000 | |
| Richmond Ballet | 2020-06-30 | 39.8449675 | |
| Atlanta Ballet | 2014-07-31 | -38.6514489 | 1 |
| Atlanta Ballet | 2019-07-31 | -36.6916144 | |
| The Sarasota Ballet | 2016-05-31 | 36.6266667 | - |
| Boston Ballet | 2014-06-30 | 36.2295633 | 2 |
| Eugene Ballet | 2021-06-30 | 35.3676500 | |

\end{table}

WITHOUT investments

```
## Spend down over time with new X axis
## Median by year
pct_change_inv_med <- pct_change_inv_ds %>%
  mutate(fiscal_year_dt = paste(fiscal_year, "-01-01", sep = ""),
         fiscal_year_dt = as.Date(fiscal_year_dt, "%Y-%m-%d")) %>%
  group_by(fiscal_year_dt, fiscal_year) %>%
  summarize(median = median(pct_change, na.rm = TRUE)) %>%
  mutate(organization_name = "Median")
## Spend Down over Time
pc_endo_ft <- pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  ggplot(aes(x = TaxPeriodEndDt, y = pct_change)) +
  geom_line(aes(group = organization_name, color = organization_name), alpha = 0.5, show.legend = FALSE) +
  scale_colour_manual(values = pal) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance, Zoomed in to 200%",
       subtitle = "Black Dot: fiscal-year median.") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  coord_cartesian(ylim=c(-100, 200)) +
  scale_x_date(breaks = scales::pretty_breaks(n = 20)) +
  geom_point(data = pct_change_inv_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 1)
pc_endo_ft
```

Percentage of Change in Endowment Balance, Zoomed in to 200%



```
pct_change_inv_ds %>%
  mutate(month = as.numeric(substring(TaxPeriodEndDt, 6, 7))) %>%
  filter(pct_change != Inf) %>%
  left_join(sp_change_years, by = c("TaxPeriodEndDt" = "date_end")) %>%
  select(organization_name, TaxPeriodEndDt, pct_change, pct_change_yr) %>%
  arrange(desc(abs(pct_change))) %>%
  make_table(title = "Percentage Change of Endowment and S&P 500",
             col_names = c("Organization Name", "Fiscal Year End Date", "Endowment Percent Change", "S&P 500 Percent Change"),
             scroll_box(height = "450px"))
```

\begin{table}

\caption{

Percentage Change of Endowment and S&P 500

}

| Organization Name | Fiscal Year End Date | Endowment Percent Change | S&P 500 Percent |
|----------------------------|----------------------|--------------------------|-----------------|
| Joffrey Ballet | 2015-06-30 | 3091.4016854 | |
| Fort Wayne Ballet | 2014-06-30 | 1803.3224138 | 2 |
| First State Ballet Theatre | 2012-12-31 | 1448.0762683 | |
| Ballet Arizona | 2016-06-30 | 586.1374894 | |
| Orlando Ballet | 2017-06-30 | 547.7003508 | 1 |
| Ballet Hispánico | 2021-06-30 | 426.7571012 | 3 |
| Nashville Ballet | 2011-05-31 | 275.0123756 | |
| First State Ballet Theatre | 2020-12-31 | 242.6848638 | |
| First State Ballet Theatre | 2020-12-31 | 242.6848638 | |
| Orlando Ballet | 2018-06-30 | 206.4099099 | |
| Atlanta Ballet | 2017-07-31 | 202.3164983 | 1 |
| Milwaukee Ballet | 2011-07-31 | 197.6406871 | |
| Nashville Ballet | 2016-05-31 | 187.7731485 | - |
| Grand Rapids Ballet | 2015-06-30 | 181.5004969 | |
| Richmond Ballet | 2016-06-30 | 175.9261036 | |
| Charlotte Ballet | 2013-06-30 | 151.2357429 | |
| Joffrey Ballet | 2019-06-30 | 133.5419677 | |
| Texas Ballet Theater | 2015-06-30 | 133.3333333 | |
| The Charleston Ballet | 2013-05-31 | 118.3461667 | |
| Dayton Ballet | 2018-06-30 | 111.2870455 | |
| San Francisco Ballet | 2012-06-30 | -101.4981454 | |
| Ballet Memphis | 2012-06-30 | 99.9715525 | |
| Ballet Austin | 2013-07-31 | 96.8010136 | |
| Ballet Memphis | 2018-06-30 | 95.9769923 | |
| Nashville Ballet | 2019-05-31 | -94.2002001 | |
| Orlando Ballet | 2020-06-30 | -91.0428942 | |
| Aspen Santa Fe Ballet | 2018-08-31 | -90.9315940 | 1 |
| Joffrey Ballet | 2020-06-30 | 86.1363009 | |
| Texas Ballet Theater | 2018-06-30 | 85.6027576 | |
| BalletMet | 2018-06-30 | 85.0249761 | |
| Atlanta Ballet | 2016-07-31 | 76.1344084 | |
| First State Ballet Theatre | 2018-12-31 | -72.7468454 | - |
| Ballet Austin | 2017-07-31 | 72.1291716 | 1 |
| First State Ballet Theatre | 2019-12-31 | 69.9900000 | 2 |
| Miami City Ballet | 2021-04-30 | 65.4636289 | 4 |
| San Francisco Ballet | 2011-06-30 | -58.6702757 | |
| The Washington Ballet | 2017-06-30 | -57.6457603 | 1 |
| The Tallahassee Ballet | 2011-07-31 | 56.0747664 | |
| First State Ballet Theatre | 2016-12-31 | 53.2556470 | |
| First State Ballet Theatre | 2014-12-31 | -52.9562948 | 1 |
| The Washington Ballet | 2019-06-30 | -51.9275920 | |
| American Repertory Ballet | 2014-06-30 | 51.6506759 | 2 |
| Richmond Ballet | 2018-06-30 | 51.2753402 | |
| First State Ballet Theatre | 2015-12-31 | -49.5289330 | - |
| Atlanta Ballet | 2012-07-31 | 47.9686134 | |
| Colorado Ballet | 2012-06-30 | -45.6020434 | |
| Dayton Ballet | 2019-06-30 | 45.1776417 | |
| Atlanta Ballet | 2013-07-31 | -44.6403251 | |
| Orlando Ballet | 2016-06-30 | -44.3305946 | |
| Richmond Ballet | 2020-06-30 | 43.9335422 | |
| Atlanta Ballet | 2014-07-31 | -43.5400727 | 1 |
| Ballet Hispánico | 2014-06-30 | 41.4922805 | 2 |
| Atlanta Ballet | 2019-07-31 | -38.7833870 | |
| The Sarasota Ballet | 2016-05-31 | 37.6855000 | - |
| American Repertory Ballet | 2013-06-30 | 34.4003625 | |
| Atlanta Ballet | 2011-07-31 | -33.2349745 | |
| Atlanta Ballet | 2020-07-31 | 32.3480884 | |

\end{table}

Rank plots about how they perform relative to the S&P 500? ## Looking at Companies who Drop at Any Point Below 40%

```
## Table for less than 40%
low_forty <- pct_change_inv_ds %>%
  filter(pct_change < -40) %>%
  select(organization_name, pct_change, BeginningYearBalanceAmt, EndYearBalanceAmt, fiscal_year)
##Table
low_forty %>%
  arrange(organization_name) %>%
  kbl(caption = "Endowment Percent Change Dropping Below 40% Of Beginning Year balance",
      col.names = c("Company Name", "Percent Change", "Beginning Balance", "End Balance", "Fiscal
        format="latex",
        booktabs=TRUE,
        escape=FALSE,
        linesep = "\\addlinespace",
        digits = 1) %>%
  kable_classic()
```

\begin{table}

\caption{Endowment Percent Change Dropping Below 40% Of Beginning Year balance}

| Company Name | Percent Change | Beginning Balance | End Balance | Fiscal Year |
|----------------------------|----------------|-------------------|-------------|-------------|
| Aspen Santa Fe Ballet | -90.9 | 6065013 | 550000 | 2018 |
| Atlanta Ballet | -43.5 | 1706513 | 1046921 | 2014 |
| Atlanta Ballet | -44.6 | 2947203 | 1706513 | 2013 |
| Colorado Ballet | -45.6 | 182437 | 100000 | 2012 |
| First State Ballet Theatre | -72.7 | 36693 | 10000 | 2018 |
| First State Ballet Theatre | -49.5 | 35876 | 18107 | 2015 |
| First State Ballet Theatre | -53.0 | 76261 | 35876 | 2014 |
| Nashville Ballet | -94.2 | 1095624 | 61350 | 2019 |
| Orlando Ballet | -91.0 | 7732855 | 696082 | 2020 |
| Orlando Ballet | -44.3 | 613186 | 338943 | 2016 |
| San Francisco Ballet | -101.5 | 1035814 | 174 | 2012 |
| San Francisco Ballet | -58.7 | 2318646 | 1035814 | 2011 |
| The Washington Ballet | -51.9 | 621423 | 310000 | 2019 |
| The Washington Ballet | -57.6 | 1212247 | 621423 | 2017 |

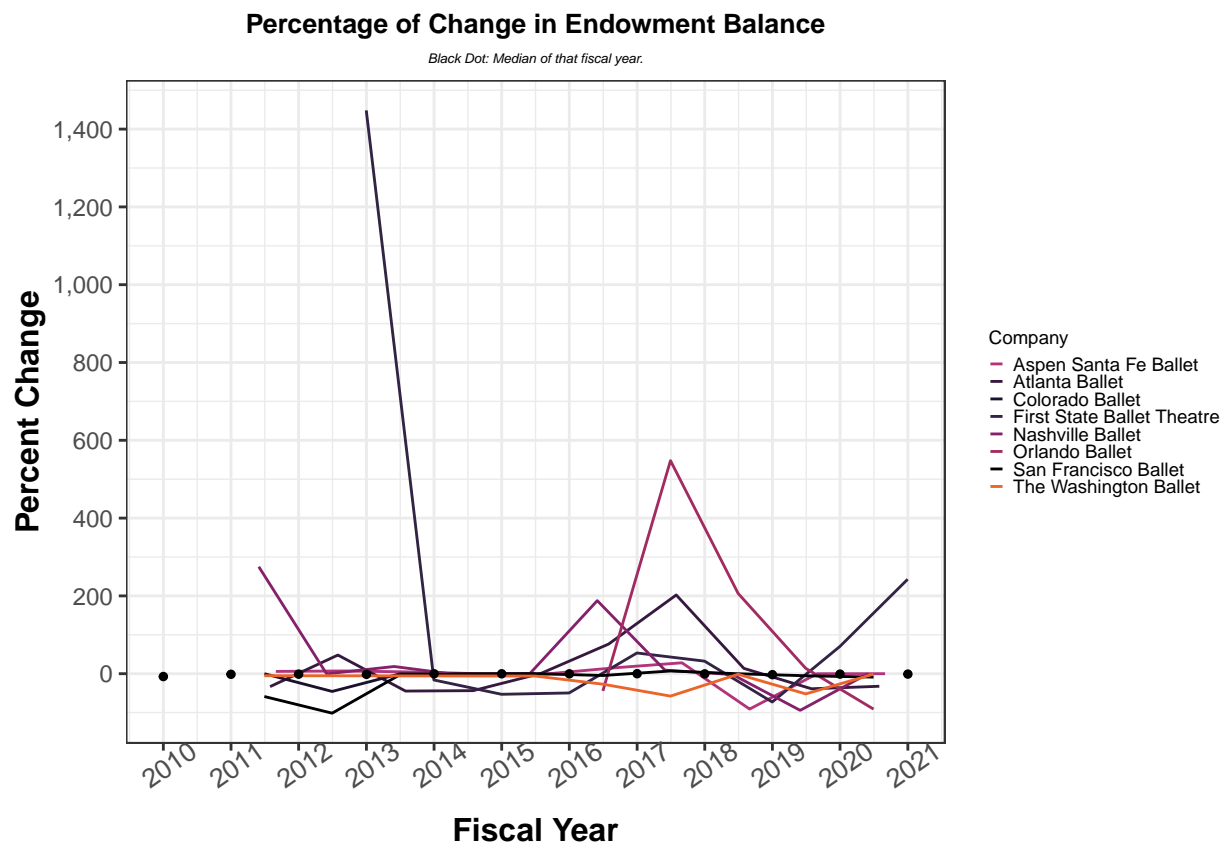
\end{table}

```
##Plotting and labeling only those with dips below 40
below_40_fr <- pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  filter(organization_name %in% low_forty$organization_name) %>%
  ggplot(aes(x = TaxPeriodEndDt, y = pct_change)) +
```

```

geom_line(aes(group = organization_name, color = organization_name), show.legend = TRUE) +
scale_colour_manual(values = pal) +
theme_bw() +
labs(y = "Percent Change",
     x = "Fiscal Year",
     title = "Percentage of Change in Endowment Balance",
     subtitle = "Black Dot: Median of that fiscal year.",
     color = "Company") +
theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
      axis.title = element_text(size = 12, face = "bold"),
      plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
      axis.text.x = element_text(size = 10, angle = 35),
      strip.text = element_text(face="bold",size = 5),
      legend.key.size = unit(2, 'mm'),
      legend.text = element_text(size=7),
      legend.title = element_text(size = 7)) +
scale_y_continuous(labels = scales::comma_format(),
                   breaks = scales::pretty_breaks(n = 10)) +
scale_x_date(breaks = scales::pretty_breaks(n = 10)) +
geom_point(data = pct_change_inv_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 1)
below_40_fr

```



Looking at Companies who Grow above 200% At Any Point

```
## Table for higher than 200%
high_200 <- pct_change_inv_ds %>%
  filter(pct_change >200) %>%
  filter(pct_change != Inf) %>%
  select(organization_name, pct_change, BeginningYearBalanceAmt, EndYearBalanceAmt, fiscal_year)
high_200 %>%
  arrange(organization_name) %>%
  kbl(caption = "Endowment Percent Change Increasing Beyond 200% Of Beginning Year balance",
      col.names = c("Company Name", "Percent Change", "Beginning Balance", "End Balance", "Fiscal Year"),
      format="latex",
      booktabs=TRUE,
      escape=FALSE,
      linesep = "\\addlinespace",
      digits = 1) %>%
  kable_classic()
```

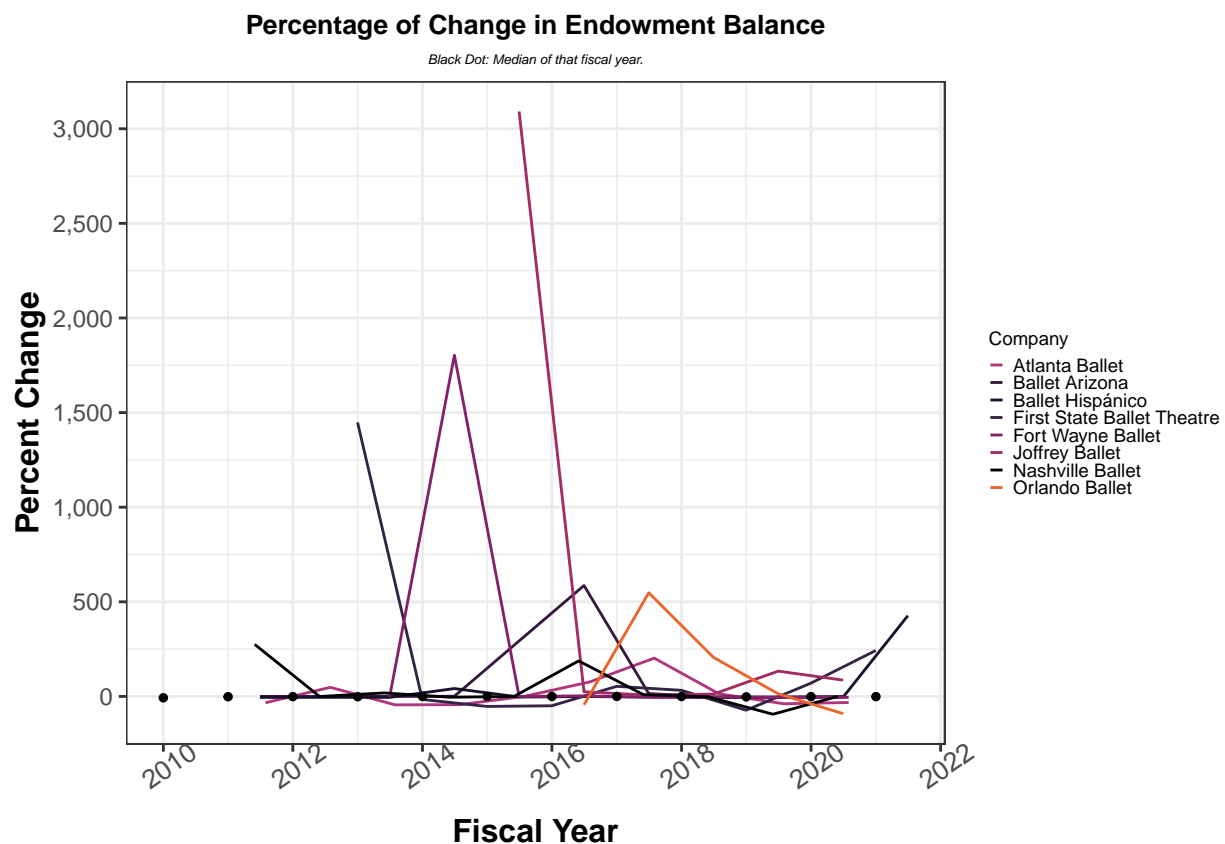
| \begin{table} | | | | |
|---|----------------|-------------------|-------------|-------------|
| \caption{Endowment Percent Change Increasing Beyond 200% Of Beginning Year balance} | | | | |
| Company Name | Percent Change | Beginning Balance | End Balance | Fiscal Year |
| Atlanta Ballet | 202.3 | 2119967 | 6523144 | 2017 |
| Ballet Arizona | 586.1 | 601399 | 4126424 | 2016 |
| Ballet Hispánico | 426.8 | 1405952 | 7481852 | 2021 |
| First State Ballet Theatre | 242.7 | 16999 | 58253 | 2020 |
| First State Ballet Theatre | 242.7 | 16999 | 58253 | 2020 |
| First State Ballet Theatre | 1448.1 | 5874 | 90934 | 2012 |
| Fort Wayne Ballet | 1803.3 | 60137 | 1201082 | 2014 |
| Joffrey Ballet | 3091.4 | 35600 | 1136139 | 2015 |
| Nashville Ballet | 275.0 | 54543 | 212030 | 2011 |
| Orlando Ballet | 206.4 | 2212808 | 6791249 | 2018 |
| Orlando Ballet | 547.7 | 338943 | 2212808 | 2017 |
| \end{table} | | | | |

```
##Plotting and labeling only those with dips below 40
above_200_fr <- pct_change_inv_ds %>%
  filter(pct_change != Inf) %>%
  filter(organization_name %in% high_200$organization_name) %>%
  ggplot(aes(x = TaxPeriodEndDt, y = pct_change)) +
  geom_line(aes(group = organization_name, color = organization_name), show.legend = TRUE) +
  scale_colour_manual(values = pal) +
  #scale_color_viridis_d(option = "D") +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
```

```

title = "Percentage of Change in Endowment Balance",
subtitle = "Black Dot: Median of that fiscal year.",
color = "Company") +
theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
axis.title = element_text(size = 12, face = "bold"),
plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
axis.text.x = element_text(size = 10, angle = 35),
strip.text = element_text(face="bold",size = 5),
legend.key.size = unit(2, 'mm'),
legend.text = element_text(size=7),
legend.title = element_text(size = 7)) +
scale_y_continuous(labels = scales::comma_format(),
breaks = scales::pretty_breaks(n = 10)) +
scale_x_date(breaks = scales::pretty_breaks(n = 10)) +
geom_point(data = pct_change_inv_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 1)
above_200_fr

```



```

##Table
pct_change_inv_ds %>%
  filter(pct_change >200) %>%
  filter(pct_change != Inf) %>%
  select(organization_name, BeginningYearBalanceAmt, EndYearBalanceAmt, pct_change, fiscal_year) %>%
  make_table_pdf(title = "c", col_names = c("a","a","a","a","a"))

```


Investments and S&P 500 specifically

Asking the question: How much of the change in endowment balance is specifically due to investments.

Calculation:

$$\frac{\sum(Investments)}{(NewestEYB - OldestBYB)} \times 100$$

```
## (INV - BYB)/BYB * 100
## Rose has notes on her choice for this calculation
endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  mutate(InvestmentEarningsOrLossesAmt = ifelse(is.na(InvestmentEarningsOrLossesAmt), 0, InvestmentEarningsOrLossesAmt))
  group_by(EIN) %>%
  summarize(inv_sum = sum(InvestmentEarningsOrLossesAmt))
```

```
## # A tibble: 46 x 2
##   EIN      inv_sum
##   <chr>    <dbl>
## 1 042312734 8411412
## 2 131882106 17043423
## 3 132584273 31904350
## 4 132642091      39
## 5 132685755 721059
## 6 132947386 101146818
## 7 210732575      0
## 8 231629970 1831526
## 9 237101094 7298768
## 10 237247909 3528
## # i 36 more rows
```

```
inv_sum <- endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  mutate(InvestmentEarningsOrLossesAmt = ifelse(is.na(InvestmentEarningsOrLossesAmt), 0, InvestmentEarningsOrLossesAmt))
  group_by(EIN) %>%
  summarize(inv_sum = sum(InvestmentEarningsOrLossesAmt))
## Earliest beginning year balance
begin <- endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  group_by(EIN) %>%
  arrange(EIN, desc(fiscal_year)) %>%
  slice_min(n=1, order_by = fiscal_year) %>%
  ungroup() %>%
  select(EIN, BeginningYearBalanceAmt, begin_bal_year = fiscal_year)
## Latest EYB
end <- endowment_data %>%
  filter(!is.na(BeginningYearBalanceAmt)) %>%
  group_by(EIN) %>%
  arrange(EIN, desc(fiscal_year)) %>%
  slice_max(n=1, order_by = fiscal_year) %>%
  ungroup() %>%
```

```

select(EIN, EndYearBalanceAmt, end_bal_year = fiscal_year)
inv_change <- inv_sum %>%
  left_join(begin, by = "EIN") %>%
  left_join(end, by = "EIN") %>%
  mutate(inv_ch_pct = inv_sum/(EndYearBalanceAmt - BeginningYearBalanceAmt))

```

Where do I go next with this? How do I compare it to the market? # Breaking Companies up by Behavior

Using only the dataset normalized for investments, as that has more consistency.

```

# Adding var for "Consistency"
## Between mean +/- SD
#inv_consist <- pct_change_inv_ds %>%
#  group_by(organization_name) %>%
#  filter(pct_change != Inf & pct_change != -Inf) %>%
#  summarize(mean = mean(pct_change),
#            sd = sd(pct_change)) %>%
#  mutate(consistent = ifelse(between(mean + sd, -7, 7) | between(mean - sd, -7, 7), TRUE, FALSE))
## Checking if any value is outside of -7:7, can do with min/max
inv_consist <- pct_change_inv_ds %>%
  group_by(organization_name) %>%
  filter(pct_change != Inf & pct_change != -Inf) %>%
  summarize(min = min(pct_change),
            max = max(pct_change),
            sd = sd(pct_change)) %>%
  mutate(consistent = ifelse(between(min + sd, -7, 7) & between(max - sd, -7, 7), TRUE, FALSE),
        #Manual inspection showed BQC should be marked as consistent but just didnt quite make it math
        consistent = ifelse(organization_name == "Ballet Quad Cities", TRUE, consistent),
        rank = rank(sd))
## Adding consistency
inv_consist_ds <- inv_consist %>%
  select(organization_name, consistent, rank) %>%
  right_join(pct_change_inv_ds, by = "organization_name") %>%
  filter(pct_change != Inf & pct_change != -Inf)

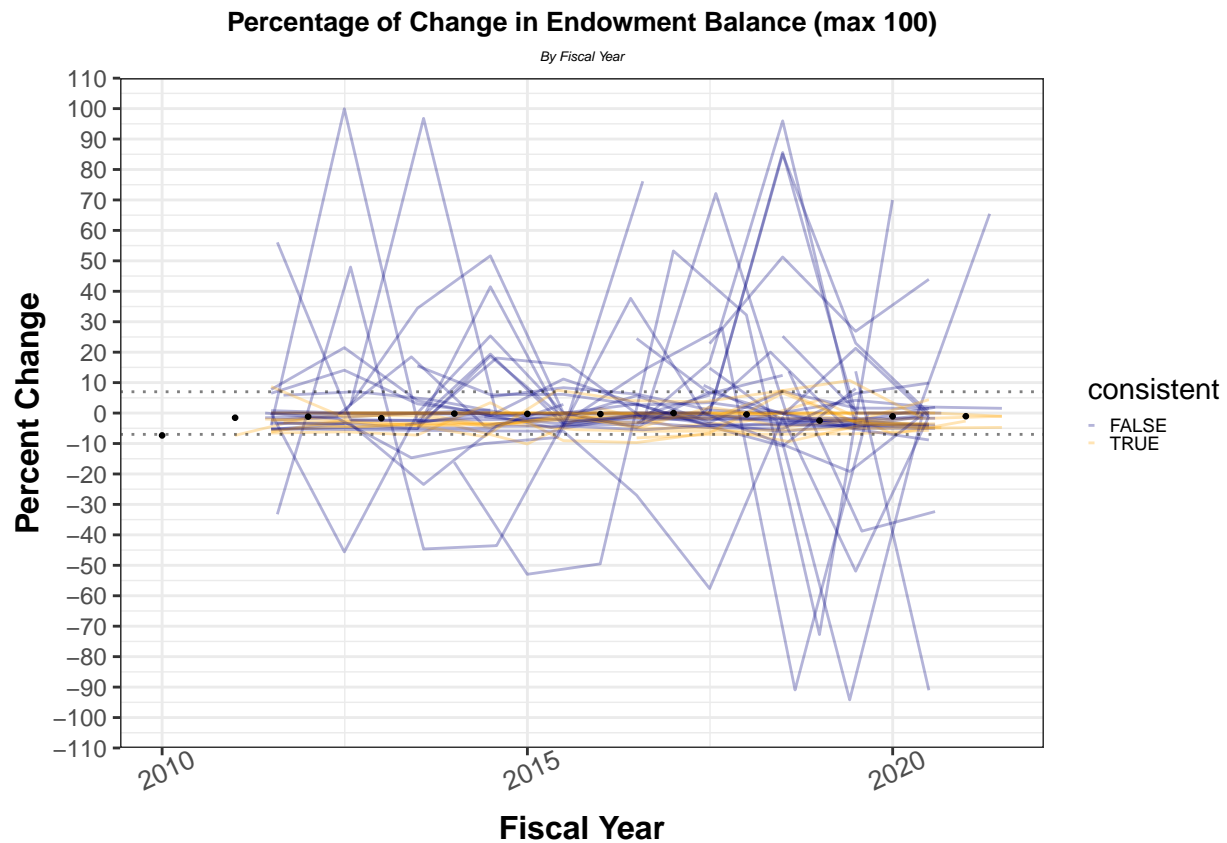
##plotting
consist_plot <- ggplot(inv_consist_ds, aes(x = TaxPeriodEndDt, y = pct_change,
                                           group = organization_name, color = consistent)) +
  geom_line(show.legend = TRUE, alpha = 0.3) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance (max 100)",
       subtitle = "By Fiscal Year") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +

```

```

scale_y_continuous(labels = scales::comma_format(),
                   breaks = scales::pretty_breaks(n = 20),
                   limits = c(-100, 100)) +
scale_color_manual(values = c("navy", "orange")) +
geom_point(data = pct_change_inv_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 0.5) +
geom_hline(yintercept = 7, alpha = .5, linetype = "dotted") +
geom_hline(yintercept = -7, alpha = .5, linetype = "dotted")
consist_plot

```



```

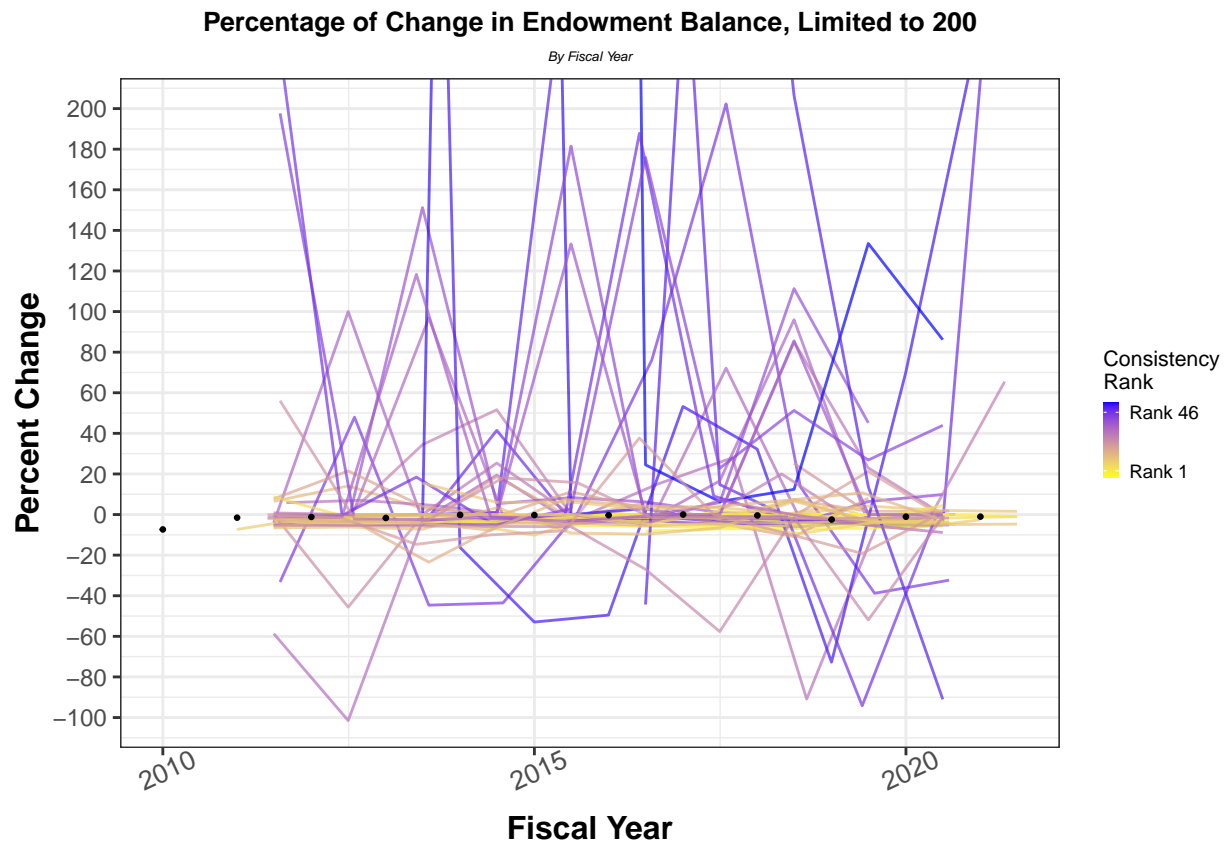
#ggplotly(consist_plot)
## Consistent by standard deviation ranking
consist_sd <- ggplot(inv_consist_ds, aes(x = TaxPeriodEndDt, y = pct_change,
                                         group = organization_name, color = rank)) +
  geom_line(show.legend = TRUE, alpha = 0.7) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance, Limited to 200",
       subtitle = "By Fiscal Year",
       color = "Consistency\nRank") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),

```

```

    legend.key.size = unit(2, 'mm'),
    legend.text = element_text(size=7),
    legend.title = element_text(size = 8)) +
  scale_y_continuous(labels = scales::comma_format(),
                     breaks = scales::pretty_breaks(n = 20)) +
  scale_color_gradient(low = "yellow", high = "blue",
                      breaks = c(5,40),
                      labels = c("Rank 1", "Rank 46")) +
  geom_point(data = pct_change_inv_med, aes(x = fiscal_year_dt, y = median), color = "black", size = 0.1) +
  coord_cartesian(ylim=c(-100, 200))
consist_sd

```



```

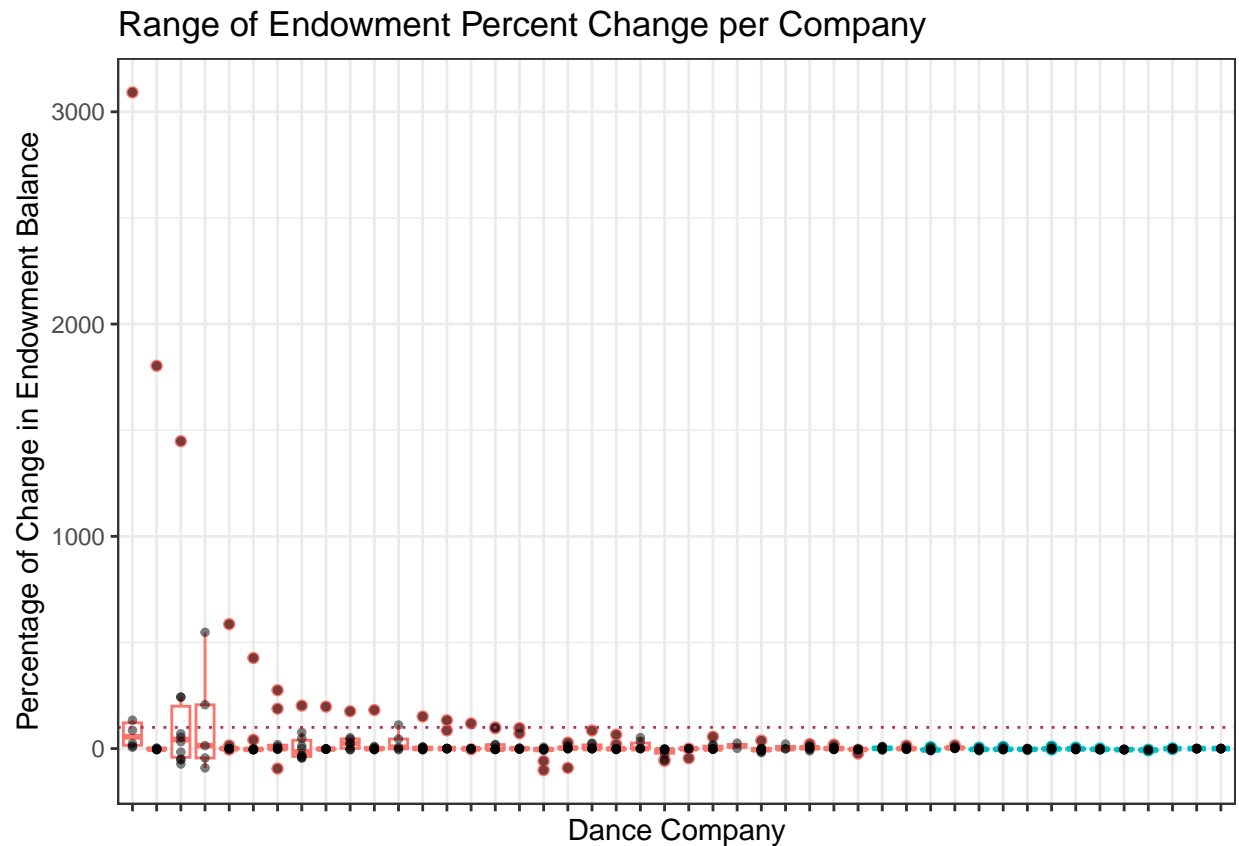
inv_box <- inv_consist_ds %>%
  group_by(organization_name) %>%
  filter(pct_change != Inf) %>%
  summarize(sd = sd(pct_change, na.rm = TRUE)) %>%
  right_join(inv_consist_ds, by = "organization_name") %>%
  select(organization_name, EIN, pct_change, sd, consistent) %>%
  mutate(organization_name = reorder(organization_name, -sd, na.rm = TRUE))
## Unlimited
box_plot <- ggplot(inv_box, aes(x = organization_name, y = pct_change)) +
  geom_boxplot(aes(color = consistent), show.legend = FALSE) +
  geom_point(size = 1, alpha = 0.5, show.legend = FALSE) +
  theme_bw() +
  labs(title = "Range of Endowment Percent Change per Company",

```

```

x = "Dance Company",
y = "Percentage of Change in Endowment Balance") +
theme(axis.text.x = element_blank()) +
geom_hline(yintercept = 100, linetype = "dotted", color = "maroon")
box_plot

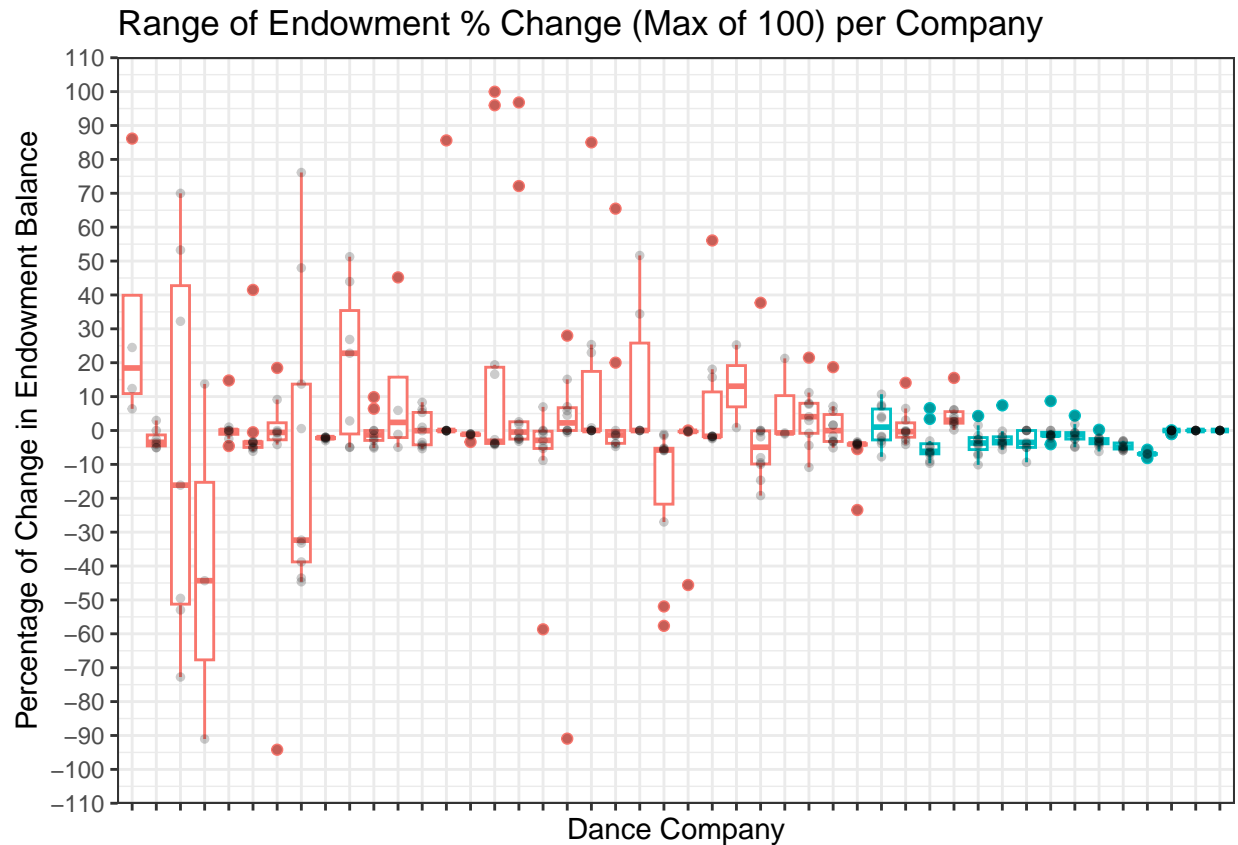
```



```

##Limited to 100 for visibility
ggplot(inv_box, aes(x = organization_name, y = pct_change)) +
  geom_boxplot(aes(color = consistent), show.legend = FALSE) +
  geom_point(size = 1, alpha = 0.2) +
  theme_bw() +
  labs(title = "Range of Endowment % Change (Max of 100) per Company",
       x = "Dance Company",
       y = "Percentage of Change in Endowment Balance") +
  theme(axis.text.x = element_blank()) +
  scale_y_continuous(breaks = scales::breaks_pretty(n = 20),
                    limit = c(-100, 100))

```



Examining the Two Classes

Lists of Companies within Each Class

```
## Names in a table
# CONSISTENT ones
inv_consist %>%
  filter(consistent == TRUE) %>%
  select(-consistent) %>%
  make_table(title = "Companies with Consistent Endowment Changes", col_names = c("Organization Name",
  scroll_box(height = "450px")
```

```
\begin{table}
\caption{
Companies with Consistent Endowment Changes
}
```

| Organization Name | Minimum Change | Maximum Change | Standard Deviation | Rank |
|---------------------------|----------------|----------------|--------------------|------|
| American Ballet Theatre | -10.1384374 | 4.2860828 | 4.0324901 | 11 |
| Ballet Quad Cities | -8.1500000 | -5.6644880 | 0.8835432 | 4 |
| Ballet West | 0.0000000 | 0.0000000 | 0.0000000 | 1 |
| Dance Theatre of Harlem | -0.0388002 | 0.0000000 | 0.0187423 | 2 |
| Eugene Ballet | -1.0345879 | 0.0688889 | 0.4221077 | 3 |
| Houston Ballet | -6.2005177 | 0.2170576 | 1.6879328 | 6 |
| Madison Ballet | -5.9572663 | -3.0421945 | 1.1388038 | 5 |
| Nevada Ballet Theatre | -4.0666759 | 8.6980699 | 3.2556331 | 8 |
| New York City Ballet | -5.6655590 | 7.4146906 | 3.6785695 | 10 |
| Pacific Northwest Ballet | -4.9507177 | 4.3953178 | 2.8264344 | 7 |
| Pittsburgh Ballet Theatre | -9.6760767 | 6.5959063 | 5.0398509 | 13 |
| The Alabama Ballet | -9.3608502 | 0.0000000 | 3.2690717 | 9 |
| Tulsa Ballet | -7.7686509 | 10.7046483 | 5.9757380 | 15 |

\end{table}

There are 13 companies which have ‘consistent’ endowment percent change.

```
## Names in a table
# INCONSISTENT ones
inv_consist %>%
  filter(consistent == FALSE) %>%
  select(-consistent) %>%
  make_table(title = "Companies with Inconsistent Endowment Changes", col_names = c("Organization Name"
  scroll_box(height = "450px"))
```

\begin{table}

\caption{

Companies with Inconsistent Endowment Changes

}

| Organization Name | Minimum Change | Maximum Change | Standard Deviation | Rank |
|------------------------------------|----------------|----------------|--------------------|------|
| Alvin Ailey American Dance Theater | -4.1243918 | 14.081909 | 5.448452 | 14 |
| American Repertory Ballet | 0.0000000 | 51.650676 | 22.878143 | 25 |
| Aspen Santa Fe Ballet | -90.9315940 | 27.985398 | 32.123542 | 28 |
| Atlanta Ballet | -44.6403251 | 202.316498 | 77.865045 | 39 |
| Ballet Arizona | -4.5855094 | 586.137489 | 195.120766 | 42 |
| Ballet Austin | -3.1962055 | 96.801014 | 38.214646 | 30 |
| Ballet Des Moines | 0.8898064 | 25.250000 | 17.225258 | 21 |
| Ballet Hispánico | -6.1313369 | 426.757101 | 129.225939 | 41 |
| Ballet Memphis | -4.2013624 | 99.971553 | 41.557976 | 31 |
| BalletMet | 0.0000000 | 85.024976 | 27.080277 | 27 |
| Boston Ballet | -5.1204623 | 18.685271 | 7.187359 | 17 |
| Charlotte Ballet | -5.3581636 | 151.235743 | 47.968535 | 34 |
| Colorado Ballet | -45.6020434 | 0.000000 | 20.295490 | 23 |
| Dayton Ballet | -4.9439121 | 111.287046 | 48.991355 | 35 |
| First State Ballet Theatre | -72.7468454 | 1448.076268 | 456.341757 | 44 |
| Fort Wayne Ballet | -5.0721240 | 1803.322414 | 638.452734 | 45 |
| Grand Rapids Ballet | -5.1685752 | 181.500497 | 57.603893 | 36 |
| Joffrey Ballet | 6.3995664 | 3091.401685 | 1241.573466 | 46 |
| Kansas City Ballet | -10.8708252 | 21.485368 | 9.369251 | 18 |
| Miami City Ballet | -4.7185994 | 65.463629 | 22.940589 | 26 |
| Milwaukee Ballet | -2.9985679 | 197.640687 | 66.634896 | 38 |
| Nashville Ballet | -94.2002001 | 275.012376 | 108.069064 | 40 |
| New Mexico Ballet Company | -1.0439065 | 21.277736 | 12.791467 | 19 |
| Oregon Ballet Theatre | 0.1654999 | 15.510390 | 4.322099 | 12 |
| Orlando Ballet | -91.0428942 | 547.700351 | 261.257026 | 43 |
| Pennsylvania Ballet | -23.4542637 | -3.339976 | 6.153578 | 16 |
| Richmond Ballet | -4.9885620 | 175.926104 | 59.163609 | 37 |
| San Francisco Ballet | -101.4981454 | 6.966102 | 34.689496 | 29 |
| Texas Ballet Theater | -0.1453333 | 133.333333 | 47.525613 | 33 |
| The Charleston Ballet | -3.3875824 | 118.346167 | 42.371447 | 32 |
| The Sarasota Ballet | -19.1999386 | 37.685500 | 15.618410 | 20 |
| The Tallahassee Ballet | -2.4204703 | 56.074766 | 18.732869 | 22 |
| The Washington Ballet | -57.6457603 | -1.092903 | 21.377074 | 24 |

\end{table}

There are 33 companies which have ‘inconsistent’ endowment percent change.

Size of Companies in Each Class

```
##Getting employee data
source(here("GET_VARS.R"))
##Specifically reading in employee data
files <- dir( here("ballet_990_released_20230208"),
             full.names = TRUE)
employ_data <- map_df(files, ~
  get_df(variables = c("//Return//ReturnData//TotalEmployeeCnt"),
          filename = .x
        )) %>%
  mutate(TotalEmployeeCnt = as.numeric(TotalEmployeeCnt)) %>%
  left_join(names, by = "EIN")
```



```
## Most recent years for each company
employ_data$fiscal_year = as.numeric(as.character(employ_data$fiscal_year))
most_recent_yrs <- employ_data %>%
  group_by(organization_name) %>%
  summarize(recent_year = max(fiscal_year, na.rm = TRUE)) %>%
  left_join(employ_data, by = c("organization_name", "recent_year" = "fiscal_year"))
## Limiting to ONLY most recent year
most_recent_ds <- inv_consist_ds %>%
  left_join(most_recent_yrs, by = "organization_name") %>%
  filter(fiscal_year == recent_year)
```

```
## Table of Consistent
most_recent_ds %>%
  filter(consistent == TRUE) %>%
  select(organization_name, TotalEmployeeCnt, EndYearBalanceAmt, recent_year) %>%
  make_table("Size of Consistent Companies, by Employees and Endowment", col_names = c("Organization Name", "Total Employee Count", "End Year Balance Amount", "Fiscal Year"))
  scroll_box(height = "450px")
```

```
\begin{table}
\caption{
}

```

| Organization Name | Total Employee Count | End Year Balance Amount | Fiscal Year |
|---------------------------|----------------------|-------------------------|-------------|
| American Ballet Theatre | 493 | 26365262 | 2020 |
| Ballet Quad Cities | 4 | 8764 | 2020 |
| Ballet West | 541 | 2127314 | 2020 |
| Dance Theatre of Harlem | 149 | 33505 | 2020 |
| Eugene Ballet | 72 | 62150 | 2021 |
| Eugene Ballet | NA | 62150 | 2021 |
| Eugene Ballet | 72 | 62150 | 2021 |
| Eugene Ballet | NA | 62150 | 2021 |
| Houston Ballet | 535 | 80123432 | 2020 |
| Madison Ballet | 41 | 932750 | 2020 |
| Nevada Ballet Theatre | 80 | 2470253 | 2021 |
| New York City Ballet | 1451 | 214442196 | 2020 |
| Pacific Northwest Ballet | 514 | 20779107 | 2020 |
| Pittsburgh Ballet Theatre | 235 | 10758728 | 2021 |
| The Alabama Ballet | 102 | 432673 | 2020 |
| Tulsa Ballet | 97 | 10614630 | 2020 |

```
\end{table}
```

```
## Table of Inconsistent
most_recent_ds %>%
  filter(consistent == FALSE) %>%
  filter(!is.na(TotalEmployeeCnt)) %>%
  select(organization_name, TotalEmployeeCnt, EndYearBalanceAmt, recent_year) %>%
  make_table("Size of Inconsistent Companies, by Employees and Endowment", col_names = c("Organization Name", "Total Employee Count", "End Year Balance Amount", "Fiscal Year"))
  scroll_box(height = "450px")
```

```
\begin{table}
```

\caption{
}

| Organization Name | Total Employee Count | End Year Balance Amount | Fiscal Year |
|------------------------------------|----------------------|-------------------------|-------------|
| Alvin Ailey American Dance Theater | 856 | 71754853 | 2020 |
| Aspen Santa Fe Ballet | 81 | 550000 | 2020 |
| Atlanta Ballet | 181 | 3356278 | 2020 |
| Ballet Arizona | 207 | 4606871 | 2020 |
| Ballet Austin | 283 | 8859387 | 2020 |
| Ballet Des Moines | 14 | 29858 | 2019 |
| Ballet Hispánico | 111 | 7481852 | 2021 |
| Ballet Memphis | 73 | 9982568 | 2020 |
| BalletMet | 209 | 535225 | 2020 |
| Boston Ballet | 664 | 18135662 | 2020 |
| Charlotte Ballet | 119 | 5670206 | 2020 |
| Dayton Ballet | 333 | 4108434 | 2019 |
| First State Ballet Theatre | 32 | 58253 | 2020 |
| First State Ballet Theatre | 32 | 58253 | 2020 |
| Grand Rapids Ballet | 63 | 2171052 | 2020 |
| Joffrey Ballet | 348 | 8999109 | 2020 |
| Kansas City Ballet | 209 | 11847916 | 2019 |
| Miami City Ballet | 333 | 2949392 | 2021 |
| Milwaukee Ballet | 179 | 588056 | 2019 |
| New Mexico Ballet Company | 1 | 28897 | 2020 |
| Oregon Ballet Theatre | 235 | 65023 | 2021 |
| Orlando Ballet | 113 | 696082 | 2020 |
| Pennsylvania Ballet | 300 | 3290362 | 2020 |
| Richmond Ballet | 240 | 2301816 | 2020 |
| San Francisco Ballet | 731 | 113923812 | 2020 |
| Texas Ballet Theater | 406 | 329418 | 2020 |
| The Sarasota Ballet | 109 | 734309 | 2020 |
| The Tallahassee Ballet | 34 | 12524 | 2020 |
| The Washington Ballet | 235 | 310000 | 2020 |

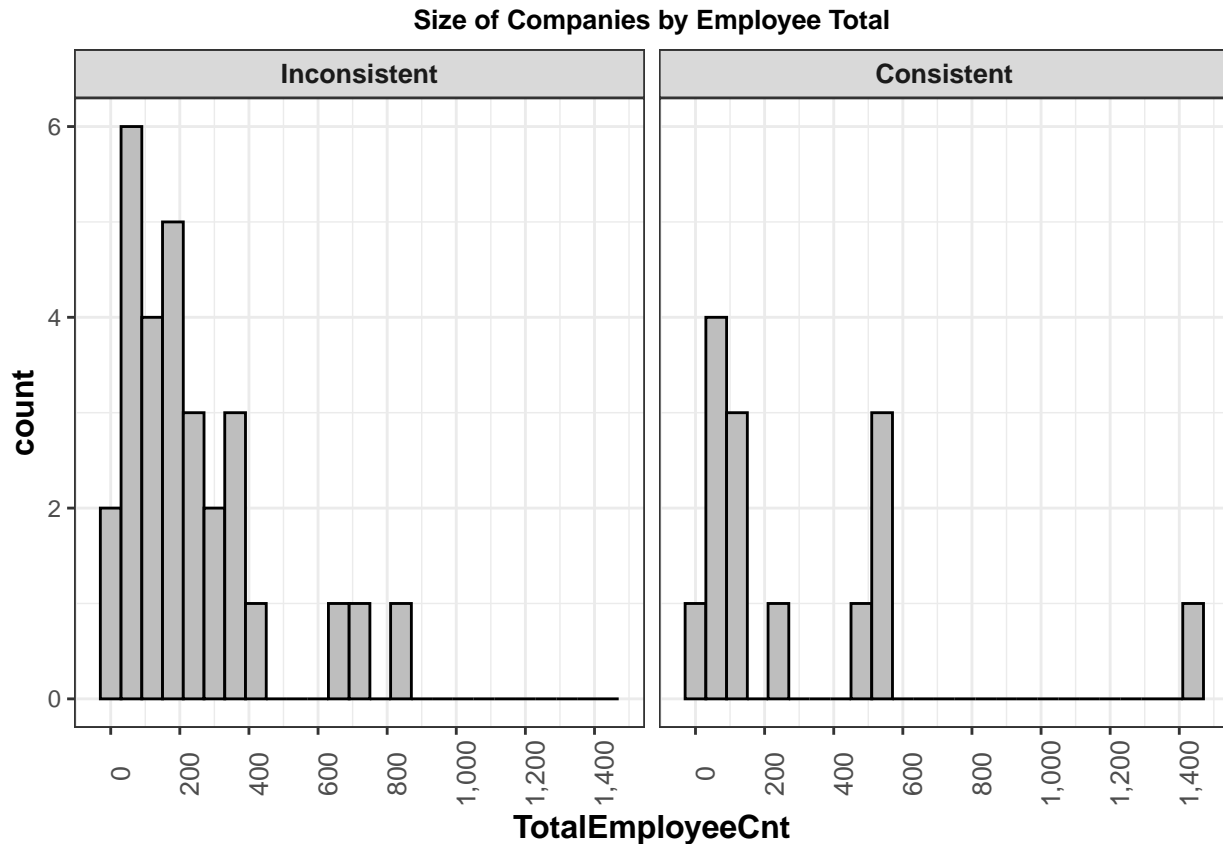
\end{table}

```
##Plotting based on consistent/inconsistent language
## Labels for facet
consistent_labels <- c(
  `TRUE` = "Consistent",
  `FALSE` = "Inconsistent"
)
##Histograms
most_recent_ds %>%
  ggplot(aes(x = TotalEmployeeCnt)) +
  geom_histogram(binwidth = 60, color = "black", fill = "gray") +
  facet_wrap(~consistent, labeller = as_labeller(consistent_labels)) +
  theme_bw() +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 90),
        strip.text = element_text(face="bold",size = 10),
```

```

legend.key.size = unit(1, 'mm'),
legend.text = element_text(size=7)) +
scale_x_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
labs( title = "Size of Companies by Employee Total")

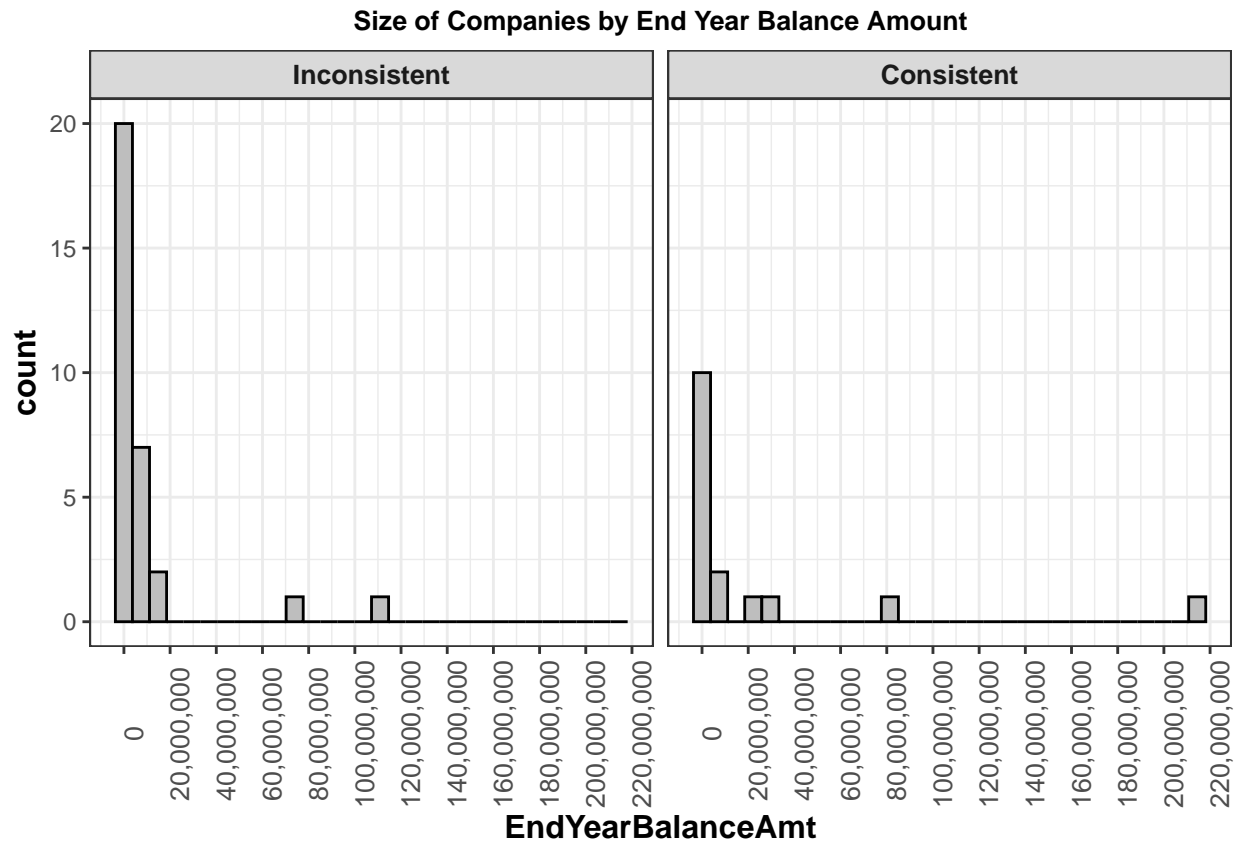
```



```

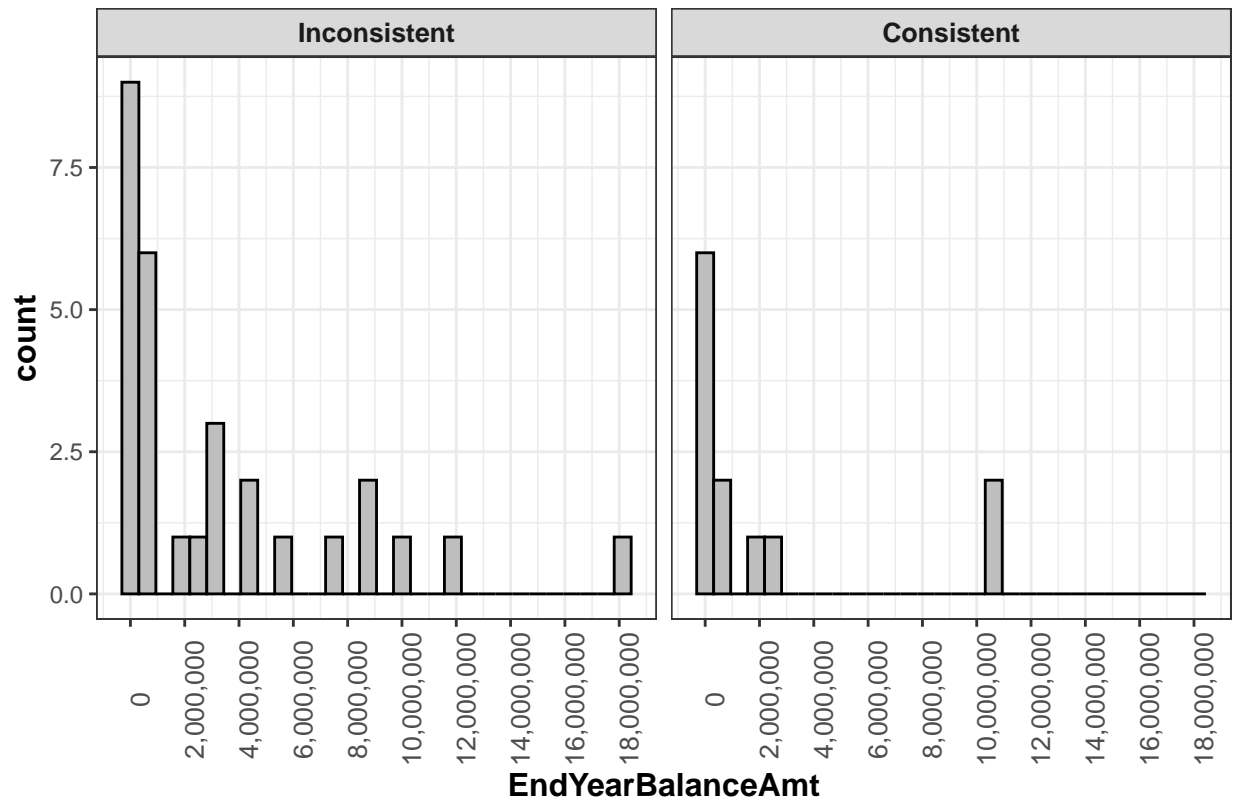
most_recent_ds %>%
  ggplot(aes(x = EndYearBalanceAmt)) +
  geom_histogram(color = "black", fill = "gray") +
  facet_wrap(~consistent, labeller = as_labeller(consistent_labels)) +
  theme_bw() +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 90),
        strip.text = element_text(face="bold",size = 10),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_x_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  labs( title = "Size of Companies by End Year Balance Amount")

```

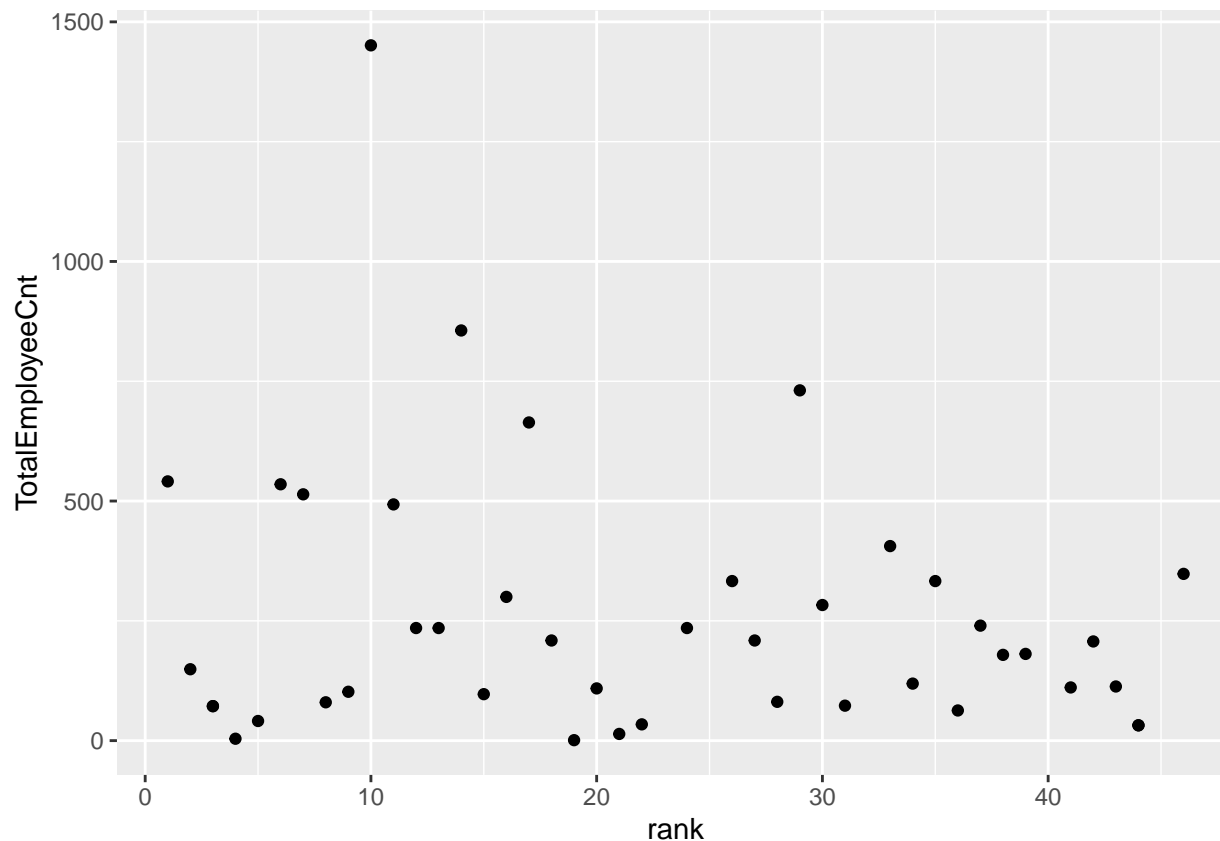


```
most_recent_ds %>%
  filter(EndYearBalanceAmt < 200000000) %>%
  ggplot(aes(x = EndYearBalanceAmt)) +
  geom_histogram(color = "black", fill = "gray") +
  facet_wrap(~consistent, labeller = as_labeller(consistent_labels)) +
  theme_bw() +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 90),
        strip.text = element_text(face="bold",size = 10),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_x_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  labs( title = "Size of Companies by End Year Balance Amount, Limit $20,000,000")
```

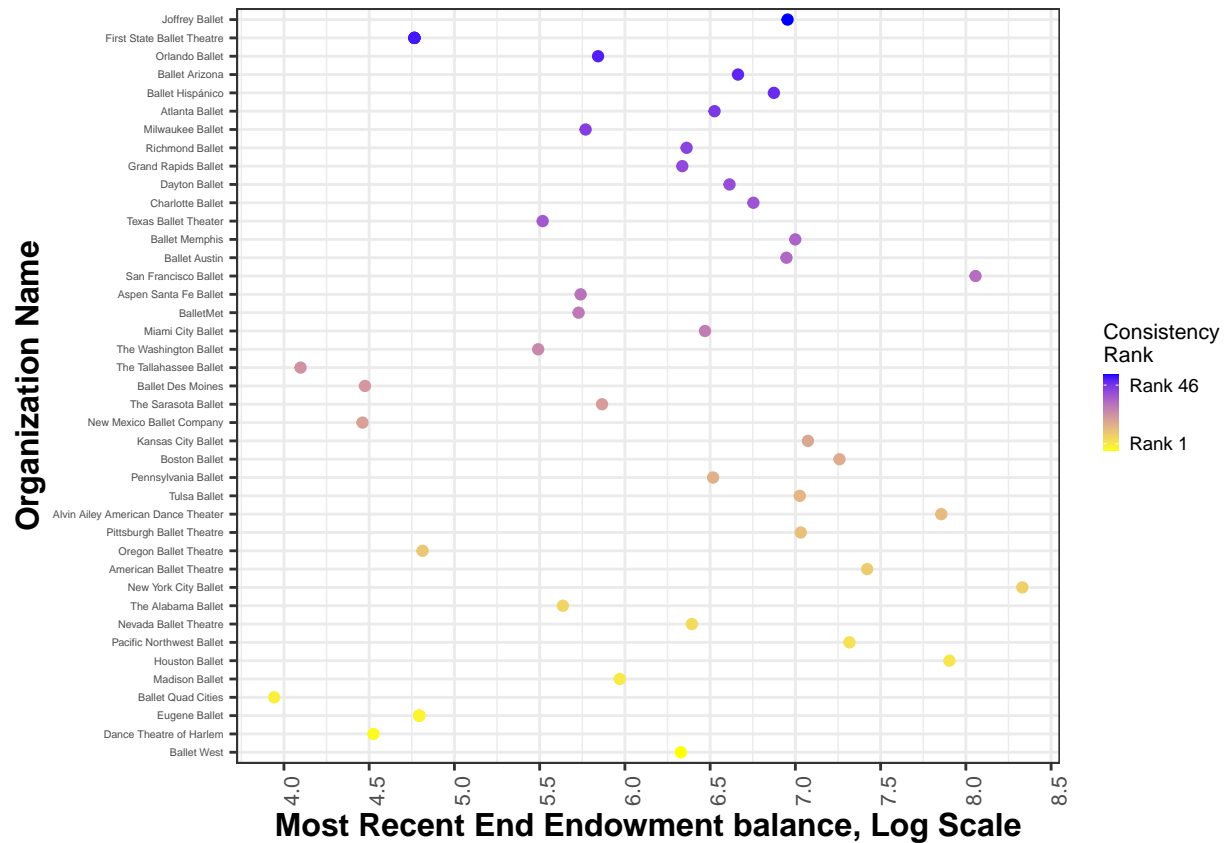
Size of Companies by End Year Balance Amount, Limit \$20,000,000



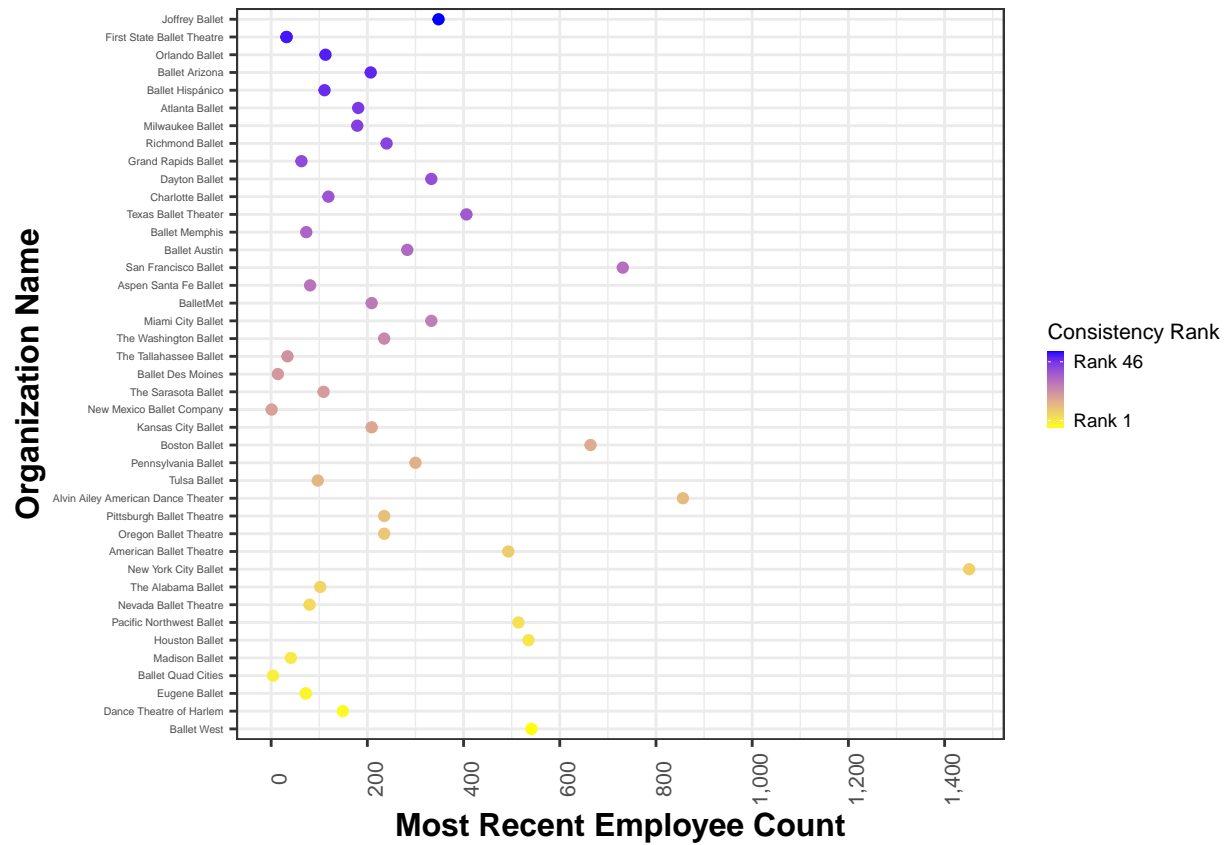
```
## Looking at size by standard deviation ranking
most_recent_ds %>%
  ggplot(aes(x = rank, y = TotalEmployeeCnt)) +
  geom_point()
```



```
##Scatter plot log scale for endowment balance
endo_size_consist_fr <- most_recent_ds %>%
  ggplot(aes(y = reorder(organization_name, rank), x = log10(EndYearBalanceAmt), color = rank)) +
  geom_point() +
  theme_bw() +
  labs(y = "Organization Name",
       x = "Most Recent End Endowment balance, Log Scale",
       color = "Consistency\nRank") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 8, angle = 90),
        axis.text.y = element_text(size = 4),
        legend.key.size = unit(2, 'mm'),
        legend.text = element_text(size=7),
        legend.title = element_text(size = 8)) +
  scale_x_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  scale_color_gradient(low = "yellow", high = "blue",
                    breaks = c(5,40),
                    labels = c("Rank 1", "Rank 46"))
endo_size_consist_fr
```



```
## Scatter plot employee count
employ_size_consist_ft <- most_recent_ds %>%
  ggplot(aes(y = reorder(organization_name, rank), x = TotalEmployeeCnt, color = rank)) +
  geom_point() +
  theme_bw() +
  labs(y = "Organization Name",
       x = "Most Recent Employee Count",
       color = "Consistency Rank") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 8, angle = 90),
        axis.text.y = element_text(size = 4),
        legend.key.size = unit(2, 'mm'),
        legend.text = element_text(size=7),
        legend.title = element_text(size = 8)) +
  scale_x_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 10)) +
  scale_color_gradient(low = "yellow", high = "blue",
                    breaks = c(5,40),
                    labels = c("Rank 1", "Rank 46"))
employ_size_consist_ft
```



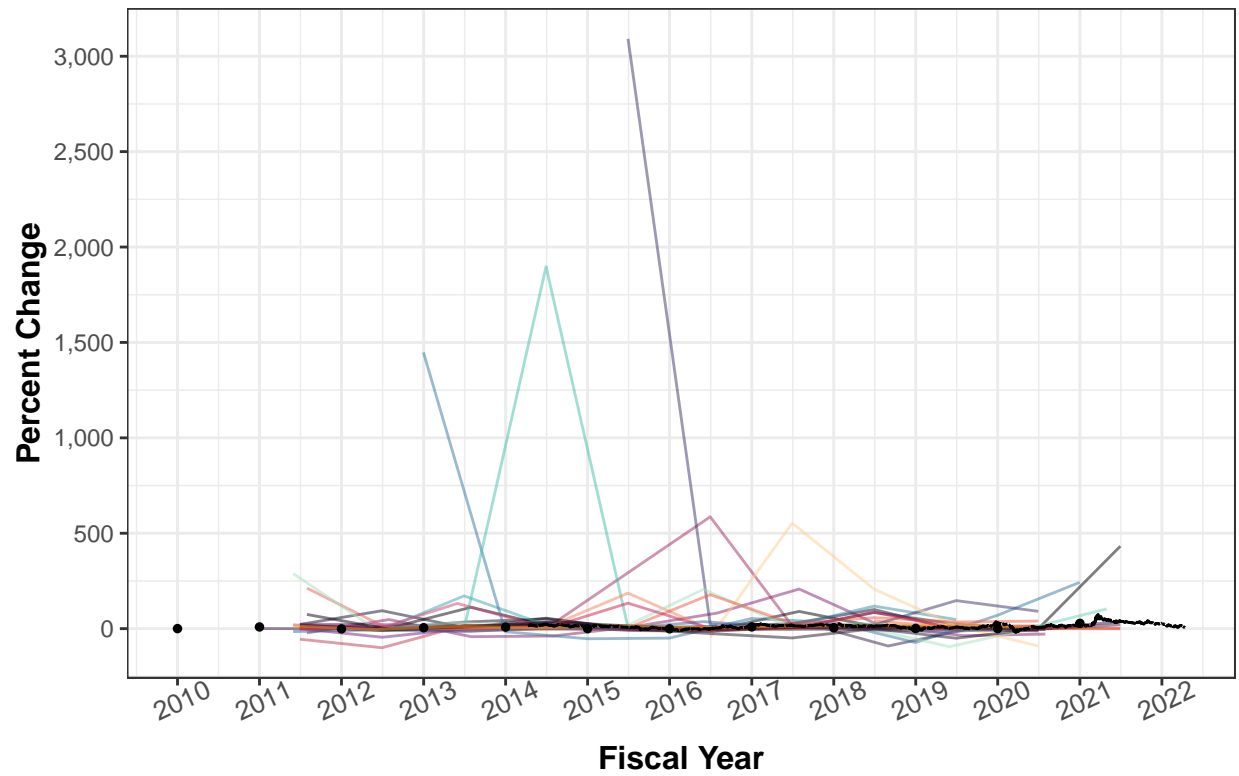
Saving Relevant Images For Report

Spend Down

pc_endow_sp_fr

Percentage of Change in Endowment Balance

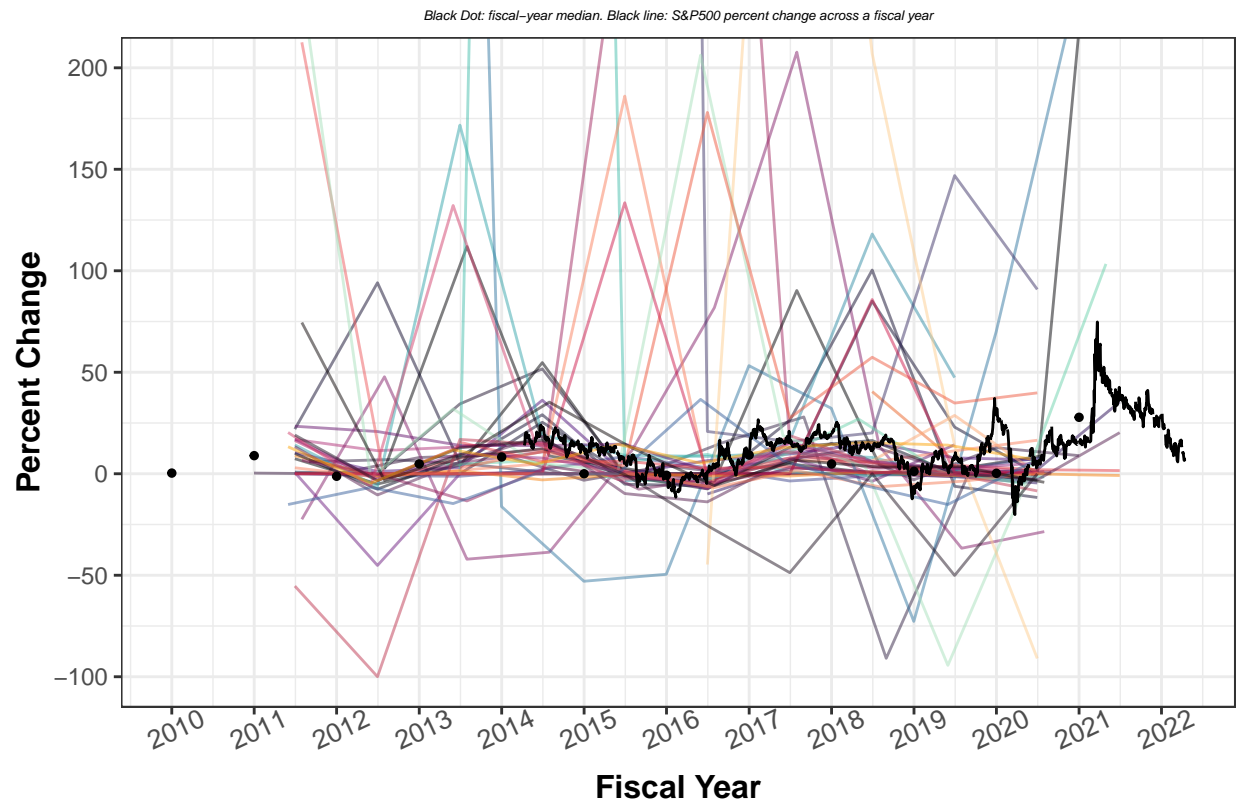
Black Dot: fiscal-year median. Black line: S&P500 percent change across a fiscal year



```
ggsave("pc_endow.png", path = here("final_report", "images"))
```

```
pc_endow_sp_fr_200
```

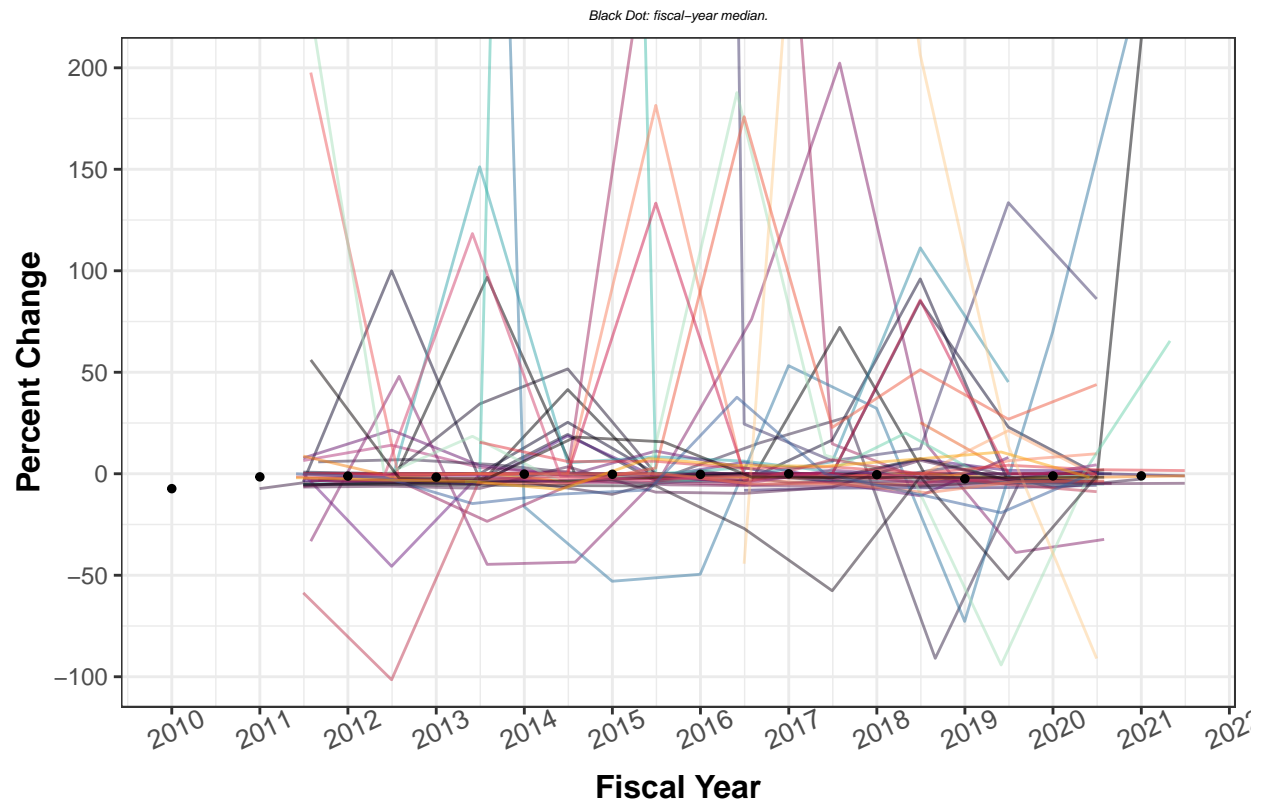
Percentage of Change in Endowment Balance, Zoomed in to 200%



```
ggsave("pc_endow_zoom.png", path = here("final_report", "images"))
```

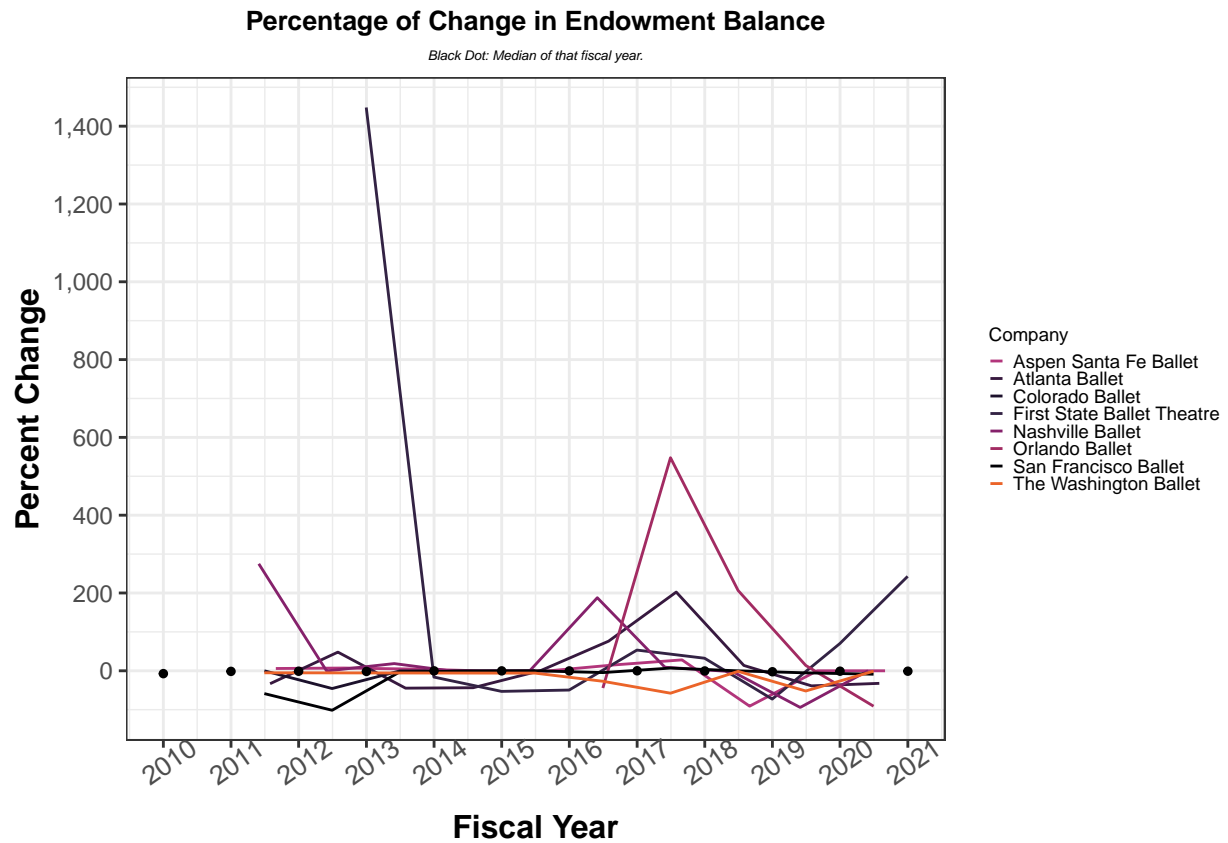
```
pc_endo_ft
```

Percentage of Change in Endowment Balance, Zoomed in to 200%



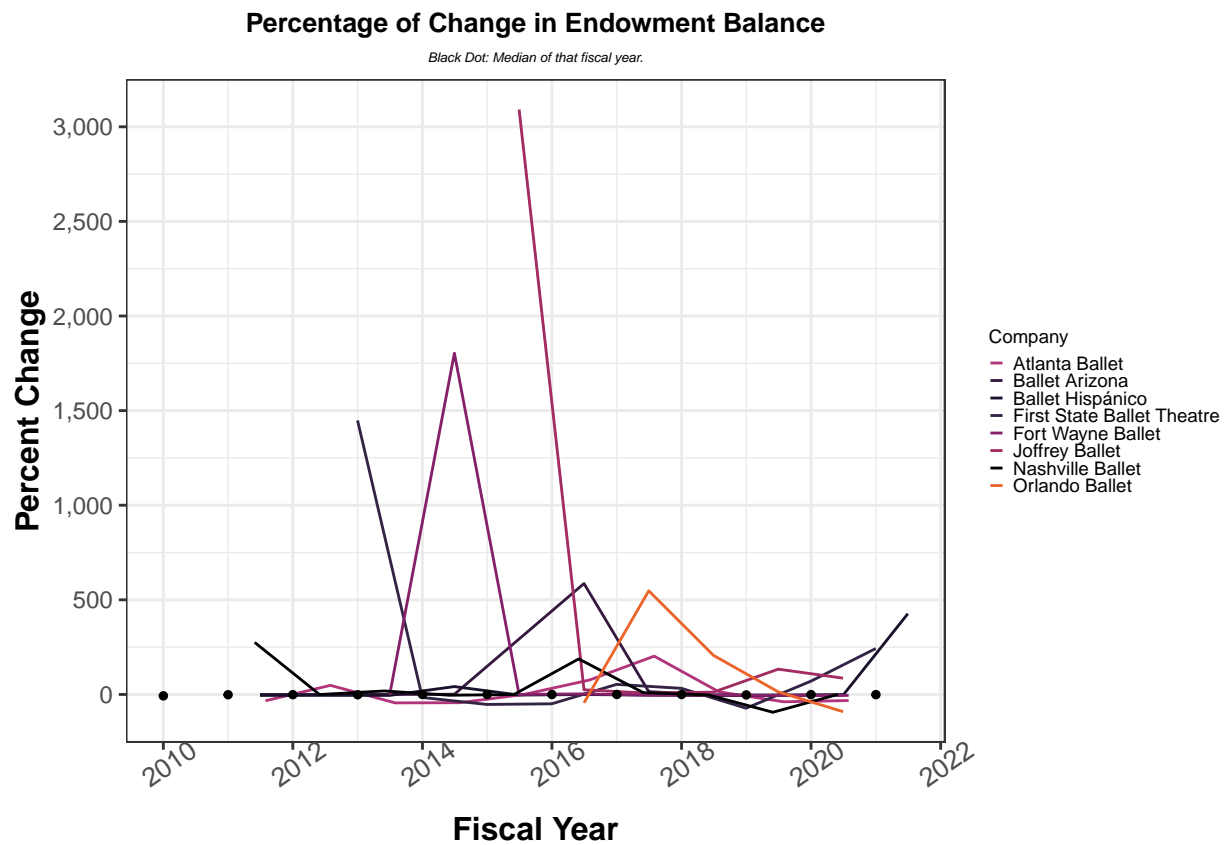
```
ggsave("pc_endow_flat.png", path = here("final_report", "images"))
```

```
below_40_fr
```



```
ggsave("pc_below_40.png", path = here("final_report", "images"))
```

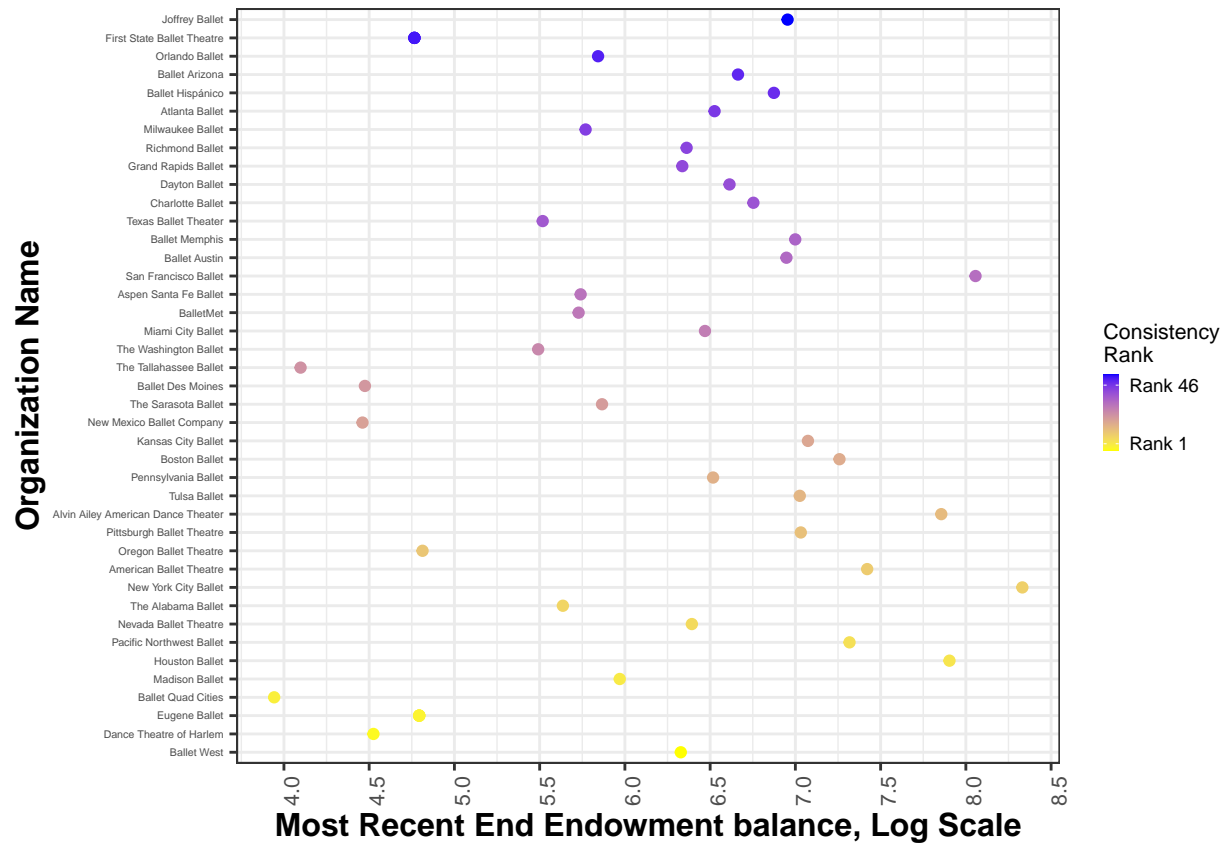
```
above_200_fr
```



```
ggsave("pc_above_200.png", path = here("final_report", "images"))
```

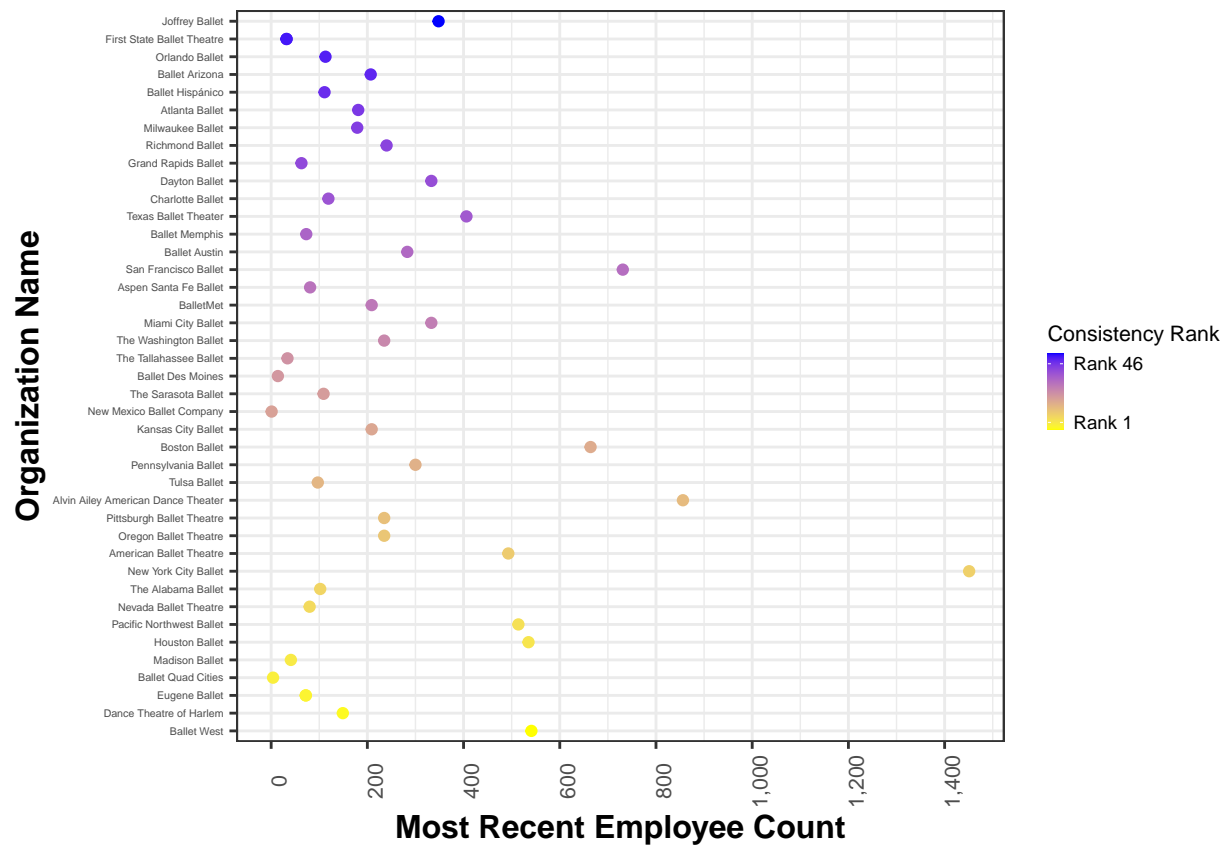
Consistency

```
endo_size_consist_fr
```



```
ggsave("consist_by_endo_size.png", path = here("final_report", "images"))
```

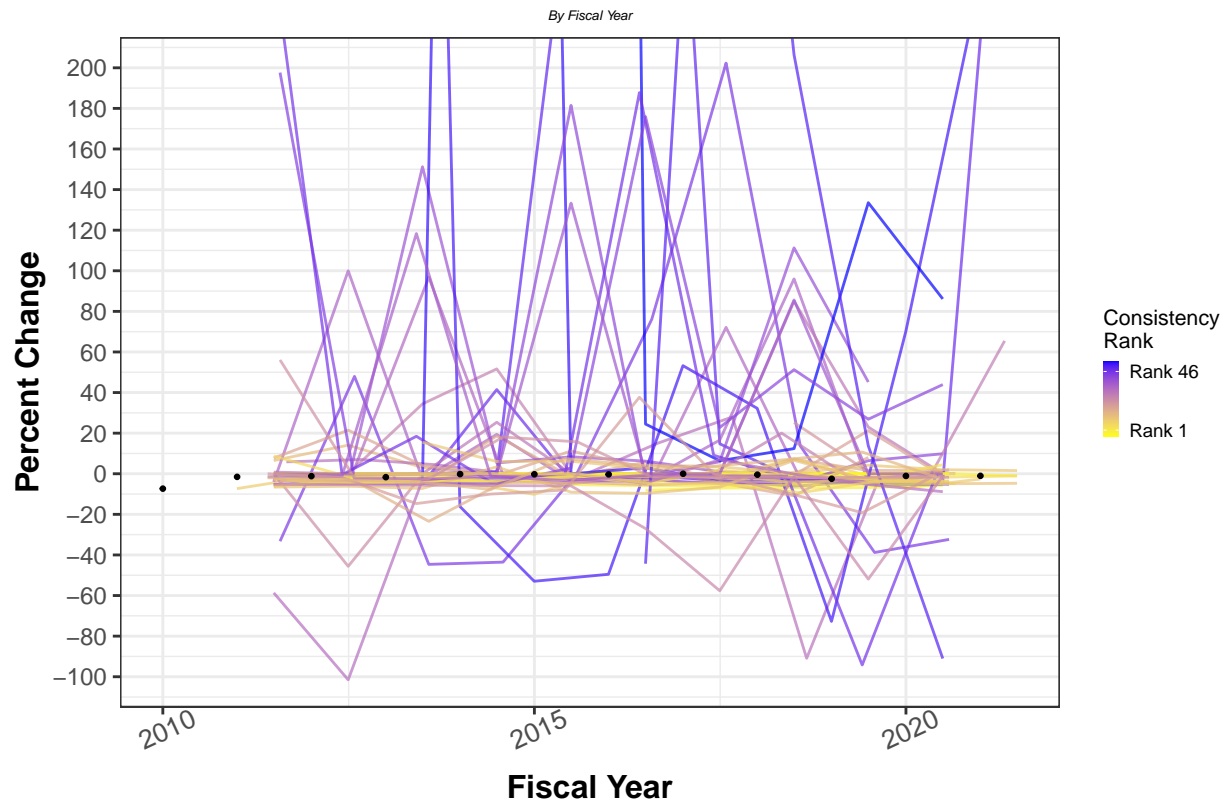
```
employ_size_consist_ft
```



```
ggsave("consist_by_employ_count.png", path = here("final_report", "images"))
```

```
consist_sd
```

Percentage of Change in Endowment Balance, Limited to 200



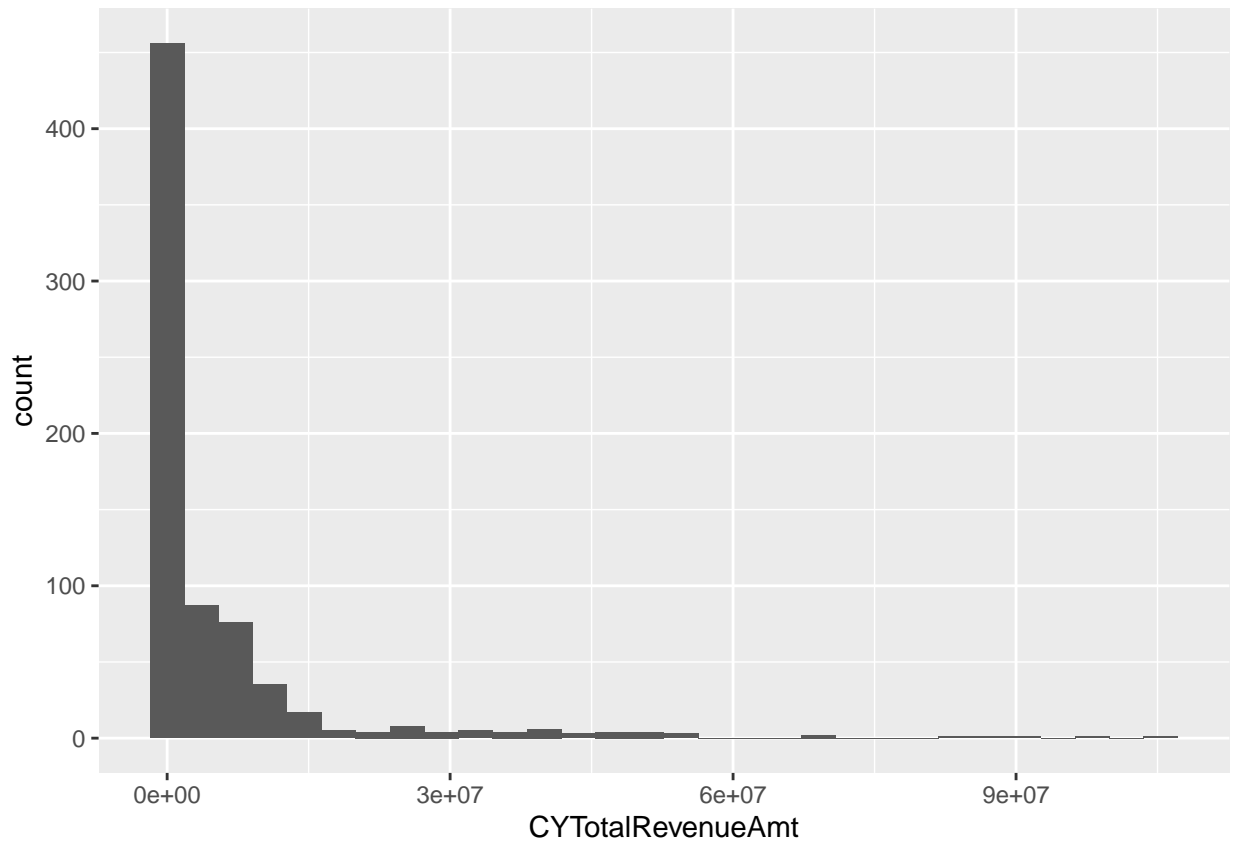
```
ggsave("pc_consist.png", path = here("final_report", "images"))
```

Total Revenue

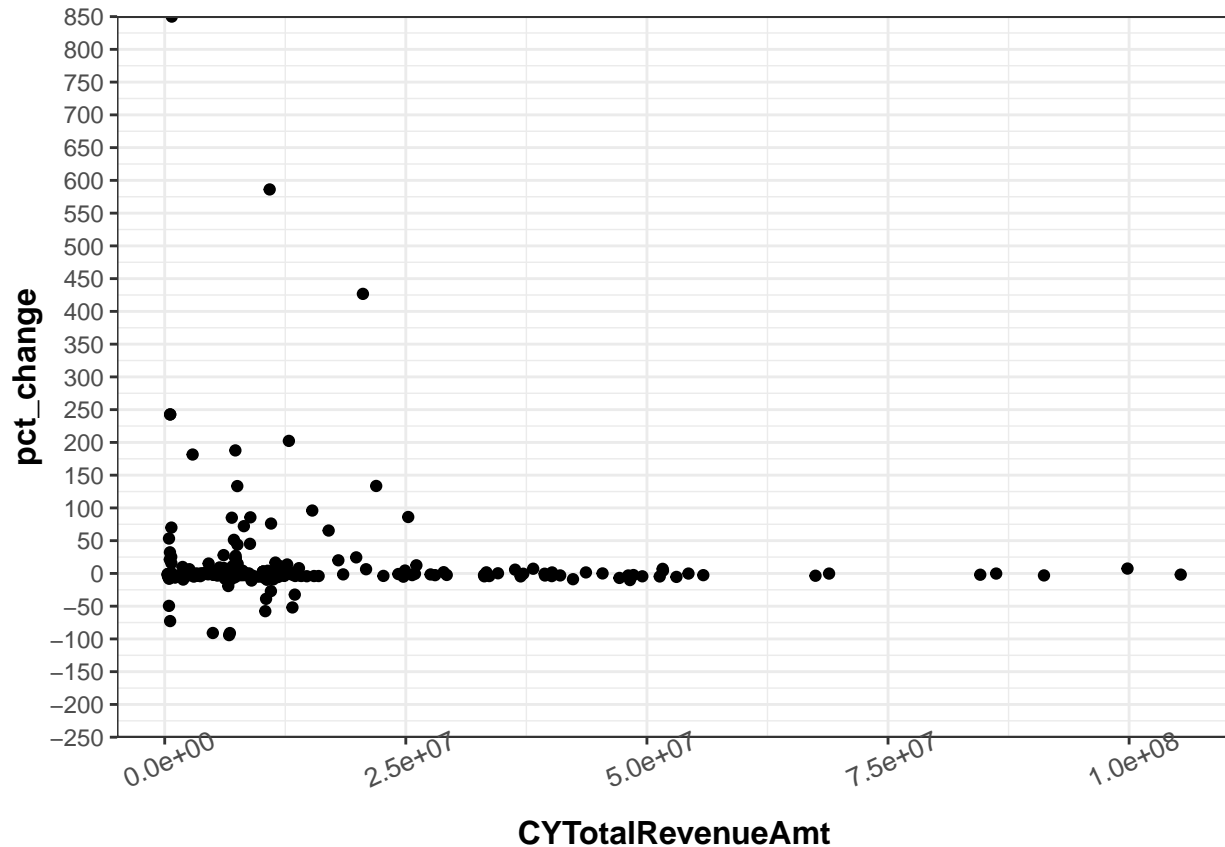
```
##Getting Total Revenue data
source(here("GET_VARS.R"))
##Specifically reading in employee data
files <- dir( here("ballet_990_released_20230208"),
              full.names = TRUE)
revenue_data <- map_df(files, ~
  get_df(variables = c("//Return//ReturnData//CYTotalRevenueAmt"),
          filename = .x
        )) %>%
  mutate(CYTotalRevenueAmt = as.numeric(CYTotalRevenueAmt)) %>%
  left_join(names, by = "EIN")
revenue_data$fiscal_year = as.numeric(as.character(revenue_data$fiscal_year))
```

Total Revenue vs Pct Change- any pattern?

```
ggplot(revenue_data, aes(x = CYTotalRevenueAmt)) +
  geom_histogram()
```

```
pct_change_inv_ds %>%
  left_join(revenue_data, by = c("EIN", "fiscal_year", "organization_name")) %>%
  ggplot(aes(x = CYTotalRevenueAmt, y = pct_change)) +
  geom_point() +
  theme_bw() +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                    breaks = scales::pretty_breaks(n = 20),
                    limits = c(-200, 800))
```



Percent Change Within the Pandemic

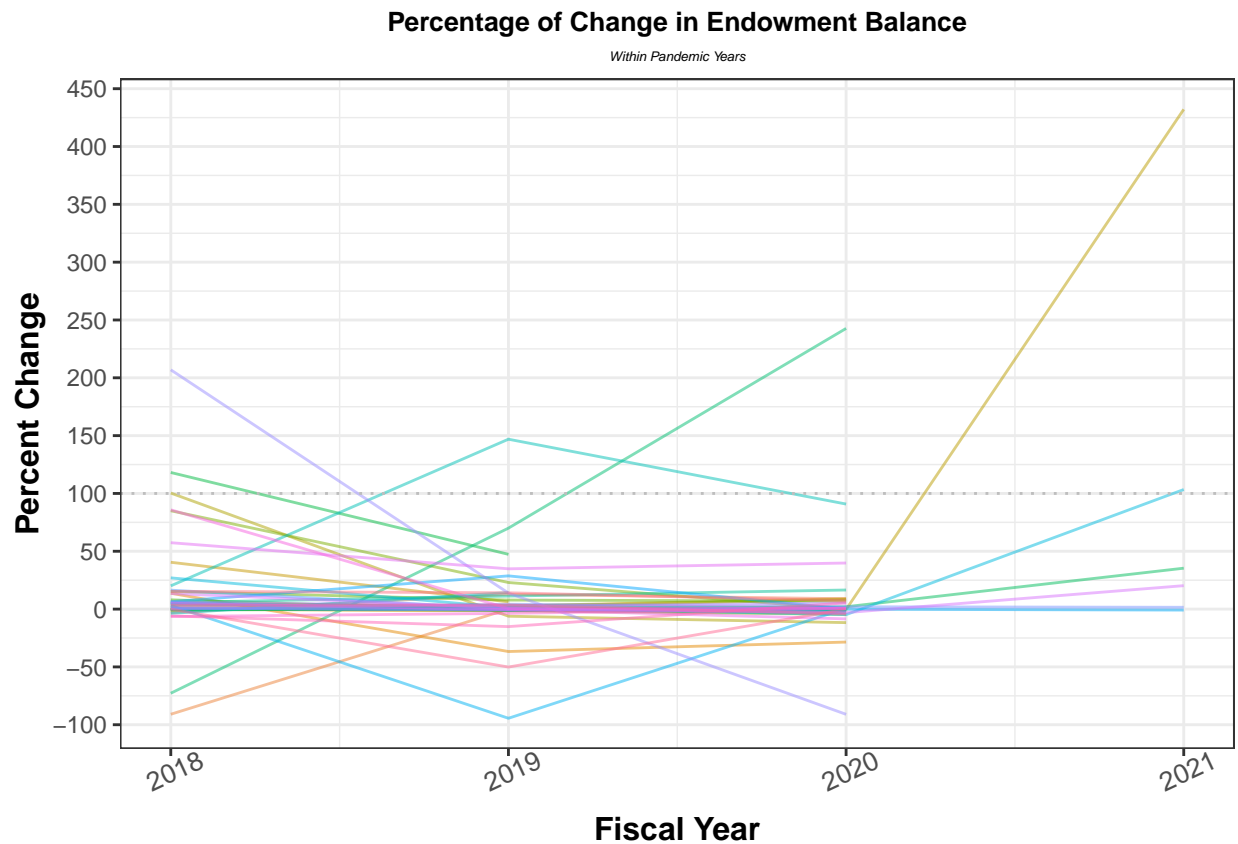
WITH investments

```
## Pandemic Years
pct_change_plot <- pct_change_ds %>%
  filter(fiscal_year %in% c("2018", "2019", "2020", "2021", "2022")) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
             group = organization_name, color = organization_name)) +
  geom_line(show.legend = FALSE, alpha = 0.5) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance",
       subtitle = "Within Pandemic Years") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
```

```

      breaks = scales::pretty_breaks(n = 20)) +
    geom_hline(yintercept = 100, linetype = "dotted", color = "gray")
pct_change_plot

```



```
#ggplotly(pct_change_plot)
```

```

## Table of available in-pandemic data
pct_change_ds %>%
  filter(fiscal_year %in% c("2019", "2020", "2021", "2022")) %>%
  select(organization_name, pct_change, fiscal_year) %>%
  arrange(desc(fiscal_year)) %>%
  make_table(title = "Percentage of Change in Endowment Balance within Pandemic Years", col_names = c("Organization", "2019", "2020", "2021", "2022"))
  scroll_box(height = "450px")

```

```

\begin{table}
\caption{
Percentage of Change in Endowment Balance within Pandemic Years
}

```

| Name | % Change | Year |
|------------------------------------|-------------|------|
| Nevada Ballet Theatre | -0.8682968 | 2021 |
| Oregon Ballet Theatre | 1.5619387 | 2021 |
| Eugene Ballet | 35.3676599 | 2021 |
| Eugene Ballet | 35.3676599 | 2021 |
| Miami City Ballet | 103.3796951 | 2021 |
| Pittsburgh Ballet Theatre | 20.2563462 | 2021 |
| Ballet Hispánico | 432.1555786 | 2021 |
| Nevada Ballet Theatre | -0.0958192 | 2020 |
| San Francisco Ballet | -8.4597794 | 2020 |
| Oregon Ballet Theatre | 1.9913020 | 2020 |
| Eugene Ballet | 2.0266667 | 2020 |
| Pacific Northwest Ballet | 5.2552006 | 2020 |
| Ballet West | 8.0887426 | 2020 |
| Ballet Arizona | -1.3712710 | 2020 |
| Texas Ballet Theater | 0.5742863 | 2020 |
| Aspen Santa Fe Ballet | 0.0000000 | 2020 |
| Ballet Austin | 8.9689362 | 2020 |
| Houston Ballet | -0.5394078 | 2020 |
| Tulsa Ballet | 7.2630153 | 2020 |
| The Sarasota Ballet | 5.3797096 | 2020 |
| The Alabama Ballet | 0.5951924 | 2020 |
| Ballet Memphis | -11.6985509 | 2020 |
| Miami City Ballet | -4.8106966 | 2020 |
| Nashville Ballet | 0.6340668 | 2020 |
| Charlotte Ballet | -3.0415463 | 2020 |
| Atlanta Ballet | -28.5158513 | 2020 |
| Richmond Ballet | 39.8449675 | 2020 |
| The Washington Ballet | 0.0000000 | 2020 |
| First State Ballet Theatre | 242.6848638 | 2020 |
| First State Ballet Theatre | 242.6848638 | 2020 |
| Ballet Quad Cities | -4.5315904 | 2020 |
| Madison Ballet | 1.0503166 | 2020 |
| Grand Rapids Ballet | 16.5116257 | 2020 |
| Joffrey Ballet | 90.8216917 | 2020 |
| Fort Wayne Ballet | -1.6147683 | 2020 |
| BalletMet | 0.0000000 | 2020 |
| Orlando Ballet | -90.9983829 | 2020 |
| The Tallahassee Ballet | -4.1994951 | 2020 |
| New Mexico Ballet Company | 0.2184921 | 2020 |
| Pittsburgh Ballet Theatre | -3.0426410 | 2020 |
| Pennsylvania Ballet | 0.9848159 | 2020 |
| New York City Ballet | -2.9450705 | 2020 |
| Ballet Hispánico | -0.2921837 | 2020 |
| Dance Theatre of Harlem | 0.0000000 | 2020 |
| Alvin Ailey American Dance Theater | 1.1173579 | 2020 |
| American Ballet Theatre | 8.9365465 | 2020 |
| Boston Ballet | 6.8383787 | 2020 |
| Ballet Des Moines | 6.2713554 | 2019 |
| Nevada Ballet Theatre | -0.9524021 | 2019 |
| San Francisco Ballet | -0.4505953 | 2019 |
| Oregon Ballet Theatre | 4.1823644 | 2019 |
| Eugene Ballet | 0.0000000 | 2019 |
| Pacific Northwest Ballet | 2.5299841 | 2019 |
| Ballet West | 2.5020351 | 2019 |
| Ballet Arizona | -1.3078878 | 2019 |
| Texas Ballet Theater | 0.5813097 | 2019 |
| Aspen Santa Fe Ballet | 0.0000000 | 2019 |

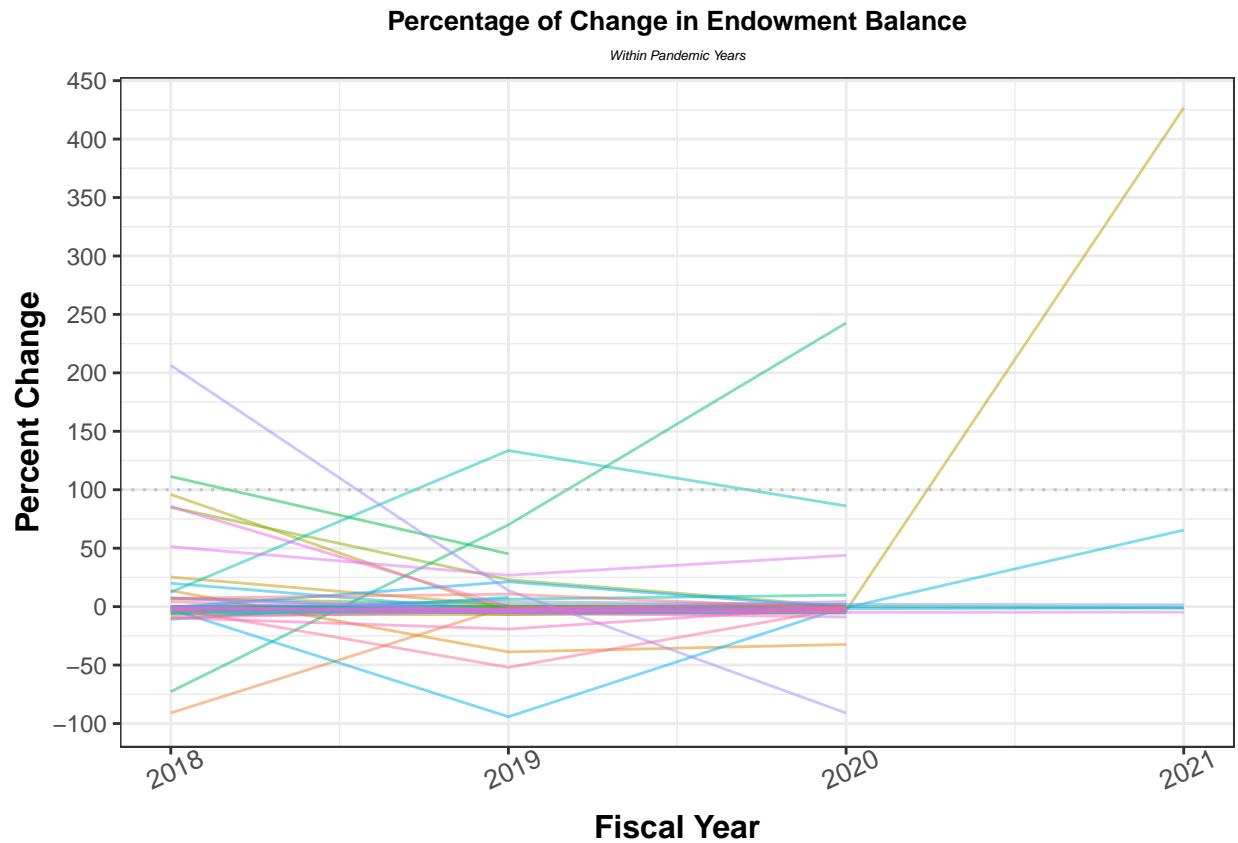
\end{table}

```
pct_change_ds %>%
  filter(fiscal_year %in% c("2019", "2020", "2021", "2022")) %>%
  select(organization_name, pct_change, fiscal_year) %>%
  group_by(fiscal_year) %>%
  summarize(total_in_year = n())
```

```
## # A tibble: 3 x 2
##   fiscal_year total_in_year
##       <dbl>         <int>
## 1      2019             43
## 2      2020             40
## 3      2021              7
```

WITHOUT investments

```
## Pandemic Years
pct_change_plot_inv <- pct_change_inv_ds %>%
  filter(fiscal_year %in% c("2018", "2019", "2020", "2021", "2022")) %>%
  ggplot(aes(x = fiscal_year, y = pct_change,
             group = organization_name, color = organization_name)) +
  geom_line(show.legend = FALSE, alpha = 0.5) +
  theme_bw() +
  labs(y = "Percent Change",
       x = "Fiscal Year",
       title = "Percentage of Change in Endowment Balance",
       subtitle = "Within Pandemic Years") +
  theme(plot.title = element_text(size = 10, face = "bold", hjust = .5),
        axis.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size = 5, face = "italic", hjust = .5),
        axis.text.x = element_text(size = 10, angle = 25),
        strip.text = element_text(face="bold",size = 5),
        legend.key.size = unit(1, 'mm'),
        legend.text = element_text(size=7)) +
  scale_y_continuous(labels = scales::comma_format(),
                     breaks = scales::pretty_breaks(n = 20)) +
  geom_hline(yintercept = 100, linetype = "dotted", color = "gray")
pct_change_plot_inv
```



```
#ggplotly(pct_change_plot_inv)
```

```
## Table of available in-pandemic data
```

```
pct_change_inv_ds %>%
```

```
  filter(fiscal_year %in% c("2019", "2020", "2021", "2022")) %>%
```

```
  select(organization_name, pct_change, fiscal_year) %>%
```

```
  arrange(desc(fiscal_year)) %>%
```

```
  make_table(title = "Percentage of Change in Endowment Balance within Pandemic Years", col_names = c("Organization", "2019", "2020", "2021", "2022"))
```

```
  scroll_box(height = "450px")
```

```
\begin{table}
```

```
\caption{
```

Percentage of Change in Endowment Balance within Pandemic Years

```
}
```

| Name | % Change | Year |
|------------------------------------|-------------|------|
| Nevada Ballet Theatre | -1.0575507 | 2021 |
| Oregon Ballet Theatre | 1.5619387 | 2021 |
| Eugene Ballet | -1.0345879 | 2021 |
| Eugene Ballet | -1.0345879 | 2021 |
| Miami City Ballet | 65.4636289 | 2021 |
| Pittsburgh Ballet Theatre | -4.7030709 | 2021 |
| Ballet Hispánico | 426.7571012 | 2021 |
| Nevada Ballet Theatre | -1.7139215 | 2020 |
| San Francisco Ballet | -8.8273535 | 2020 |
| Oregon Ballet Theatre | 1.9913020 | 2020 |
| Eugene Ballet | 0.0688889 | 2020 |
| Pacific Northwest Ballet | 4.3953178 | 2020 |
| Ballet West | 0.0000000 | 2020 |
| Ballet Arizona | -0.6848327 | 2020 |
| Texas Ballet Theater | 0.0000000 | 2020 |
| Aspen Santa Fe Ballet | 0.0000000 | 2020 |
| Ballet Austin | 1.8896472 | 2020 |
| Houston Ballet | -4.0352545 | 2020 |
| Tulsa Ballet | -1.7067692 | 2020 |
| The Sarasota Ballet | -0.4168927 | 2020 |
| The Alabama Ballet | -3.5193077 | 2020 |
| Ballet Memphis | -3.8335615 | 2020 |
| Miami City Ballet | -0.9255126 | 2020 |
| Nashville Ballet | 0.0000000 | 2020 |
| Charlotte Ballet | -4.6291277 | 2020 |
| Atlanta Ballet | -32.3489884 | 2020 |
| Richmond Ballet | 43.9335422 | 2020 |
| The Washington Ballet | -1.0929032 | 2020 |
| First State Ballet Theatre | 242.6848638 | 2020 |
| First State Ballet Theatre | 242.6848638 | 2020 |
| Ballet Quad Cities | -5.6644880 | 2020 |
| Madison Ballet | -4.7836803 | 2020 |
| Grand Rapids Ballet | 9.8477067 | 2020 |
| Joffrey Ballet | 86.1363009 | 2020 |
| Fort Wayne Ballet | -5.0721240 | 2020 |
| BalletMet | 0.0000000 | 2020 |
| Orlando Ballet | -91.0428942 | 2020 |
| The Tallahassee Ballet | -1.7058059 | 2020 |
| New Mexico Ballet Company | -1.0439065 | 2020 |
| Pittsburgh Ballet Theatre | -4.9093733 | 2020 |
| Pennsylvania Ballet | -3.8933190 | 2020 |
| New York City Ballet | -3.5223075 | 2020 |
| Ballet Hispánico | -3.6378284 | 2020 |
| Dance Theatre of Harlem | 0.0000000 | 2020 |
| Alvin Ailey American Dance Theater | -4.1243918 | 2020 |
| American Ballet Theatre | -2.6235825 | 2020 |
| Boston Ballet | 1.5323016 | 2020 |
| Ballet Des Moines | 0.8898064 | 2019 |
| Nevada Ballet Theatre | -2.2385659 | 2019 |
| San Francisco Ballet | -5.3065465 | 2019 |
| Oregon Ballet Theatre | 4.1823644 | 2019 |
| Eugene Ballet | 0.0000000 | 2019 |
| Pacific Northwest Ballet | -2.5450559 | 2019 |
| Ballet West | 0.0000000 | 2019 |
| Ballet Arizona | -4.5855094 | 2019 |
| Texas Ballet Theater | 0.0000000 | 2019 |
| Aspen Santa Fe Ballet | 0.0000000 | 2019 |

\end{table}

```
pct_change_inv_ds %>%  
  filter(fiscal_year %in% c("2019", "2020", "2021", "2022")) %>%  
  select(organization_name, pct_change, fiscal_year) %>%  
  group_by(fiscal_year) %>%  
  summarize(total_in_year = n())
```

```
## # A tibble: 3 x 2  
##   fiscal_year total_in_year  
##       <dbl>         <int>  
## 1      2019             43  
## 2      2020             40  
## 3      2021              7
```