BSA Spring 2025

2025-05-19

```
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(tidyr)
library(scales)

df <- read_excel("~/Desktop/BSA/NCAA Data Collection.xlsx", sheet = "Sheet1")
```

```
## New names:
## * `` -> `...9`
## * `` -> `...10`
```

```
years <- c("2017-2018", "2018-2019", "2019-2020", "2020-2021", "2021-2022", "2022-2023", "2023-2024")


# Average Spending

# Convert year columns to numeric
for (col in years) {
  df[[col]] <- as.numeric(df[[col]])
}
```

```
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
```

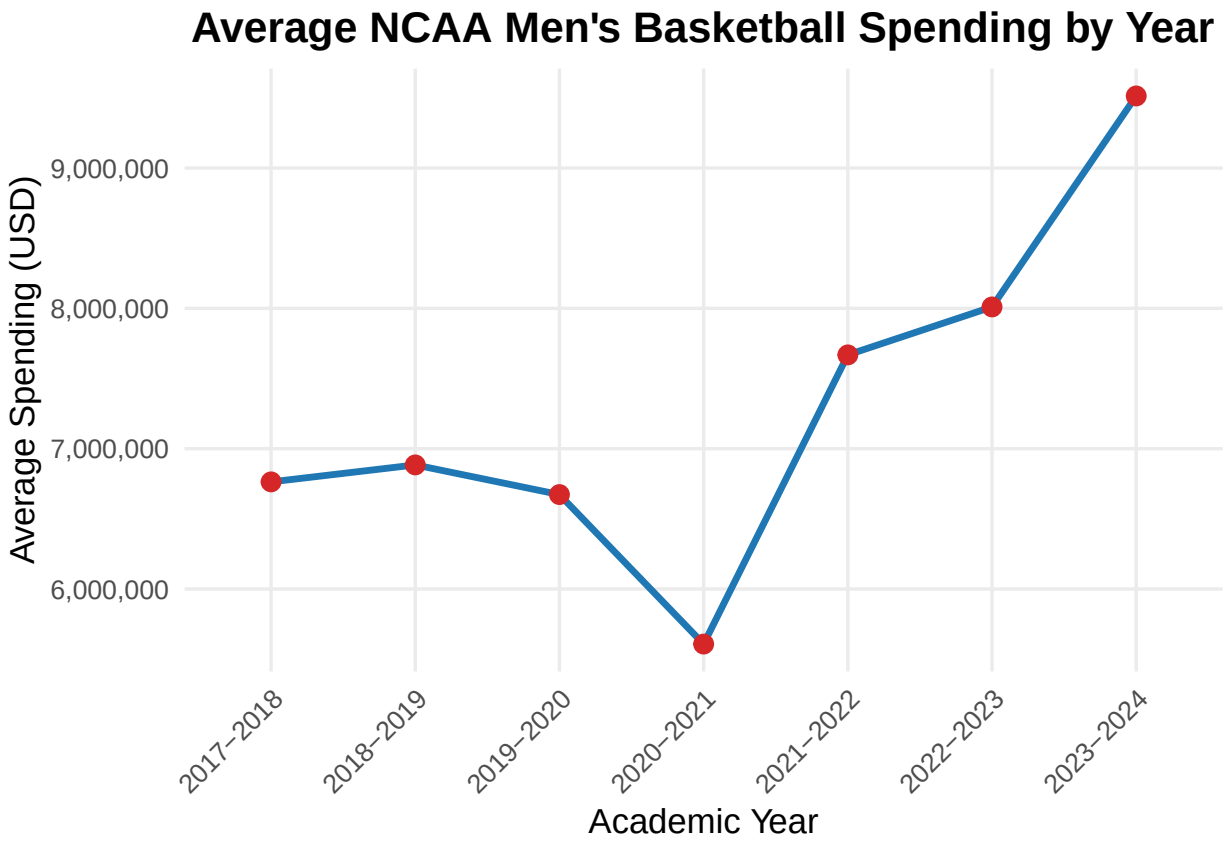
```r
# Now extract the spending data
spending <- df[ , years]

# Calculate average spending
avg_spending <- colMeans(spending, na.rm = TRUE)

# Create data frame
avg_spending_df <- data.frame(
  Year = names(avg_spending),
  Avg_Spending = as.numeric(avg_spending)
)

# Plot
ggplot(avg_spending_df, aes(x = Year, y = Avg_Spending, group = 1)) +
  geom_line(color = "#1f77b4", size = 1.2) +
  geom_point(color = "#d62728", size = 3) +
  labs(
    title = "Average NCAA Men's Basketball Spending by Year",
    x = "Academic Year",
    y = "Average Spending (USD)"
  ) +
  scale_y_continuous(labels = comma) +  # <- removes scientific notation
  theme_minimal(base_size = 13) +
  theme(
    plot.title = element_text(face = "bold", size = 16, hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1),
    panel.grid.minor = element_blank()
  )
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```




```r
# Percent Change

# Step 1: Calculate percent change manually
percent_change <- diff(avg_spending) / head(avg_spending, -1) * 100

# Step 2: Create year-to-year labels
change_labels <- paste0(
  names(avg_spending)[1:(length(avg_spending) - 1)],
  " to ",
  names(avg_spending)[2:length(avg_spending)]
)

# Step 3: Data frame for plotting
pct_change_df <- data.frame(
  Change_Label = change_labels,
  Percent_Change = percent_change
)

# Step 4: Plot with ggplot2
ggplot(pct_change_df, aes(x = Change_Label, y = Percent_Change, group = 1)) +
  geom_line(color = "#ff7f0e", size = 1.2) +
  geom_point(color = "#2ca02c", size = 3) +
  labs(
    title = "Percent Change in NCAA Men's Basketball Spending",
    x = "Year Range",
    y = "Percent Change (%)"
```

```
) +
scale_y_continuous(labels = scales::percent_format(scale = 1)) +  # Keep raw percent
theme_minimal(base_size = 13) +
theme(
  plot.title = element_text(face = "bold", size = 16, hjust = 0.5),
  axis.text.x = element_text(angle = 45, hjust = 1),
  panel.grid.minor = element_blank()
)
```