The Story of America's Dam Infrastructure

Exploring 92,428 Dams Across the United States

Executive Summary

This analysis explores the **National Inventory of Dams (NID)** dataset, containing detailed information about **92,428** dams across the United States. Our findings reveal fascinating patterns about America's water infrastructure, from the post-war construction boom to surprising geographic distributions.

> **!** Key Findings at a Glance
>
> - **Peak construction era**: 1960s saw 18,599 dams built
> - **Recreation dominates**: 30,815 dams are primarily for recreation
> - **Safety concerns**: 16,843 dams classified as **high hazard**
> - **Private ownership**: 65% of dams are privately owned

The Great Dam Building Boom

When Were America's Dams Built?

The story of American dam construction is essentially the story of **post-World War II infrastructure development**. The median dam was completed in **1965**, highlighting the massive construction efforts of the mid-20th century.

```
construction_data <- dat |>
  filter(!is.na(year_completed), year_completed > 1800) |>
  mutate(decade = floor(year_completed / 10) * 10) |>
  count(decade) |>
  mutate(
```

```r
    peak_decade = decade == 1960,
    label = if_else(peak_decade, paste0(format(n, big.mark = ","), " dams"), "")
  )

ggplot(construction_data, aes(x = decade, y = n)) +
  geom_col(aes(fill = peak_decade), show.legend = FALSE) +
  geom_text(aes(label = label), vjust = -0.5, color = "darkred", fontface = "bold") +
  scale_fill_manual(values = c("FALSE" = "steelblue", "TRUE" = "darkred")) +
  scale_y_continuous(labels = comma_format()) +
  labs(
    title = "The 1960s: America\'s Dam Building Golden Age",
    subtitle = "Nearly 19,000 dams built in a single decade",
    x = "Decade",
    y = "Number of Dams Built",
    caption = "Source: National Inventory of Dams"
  ) +
  theme_dam
```

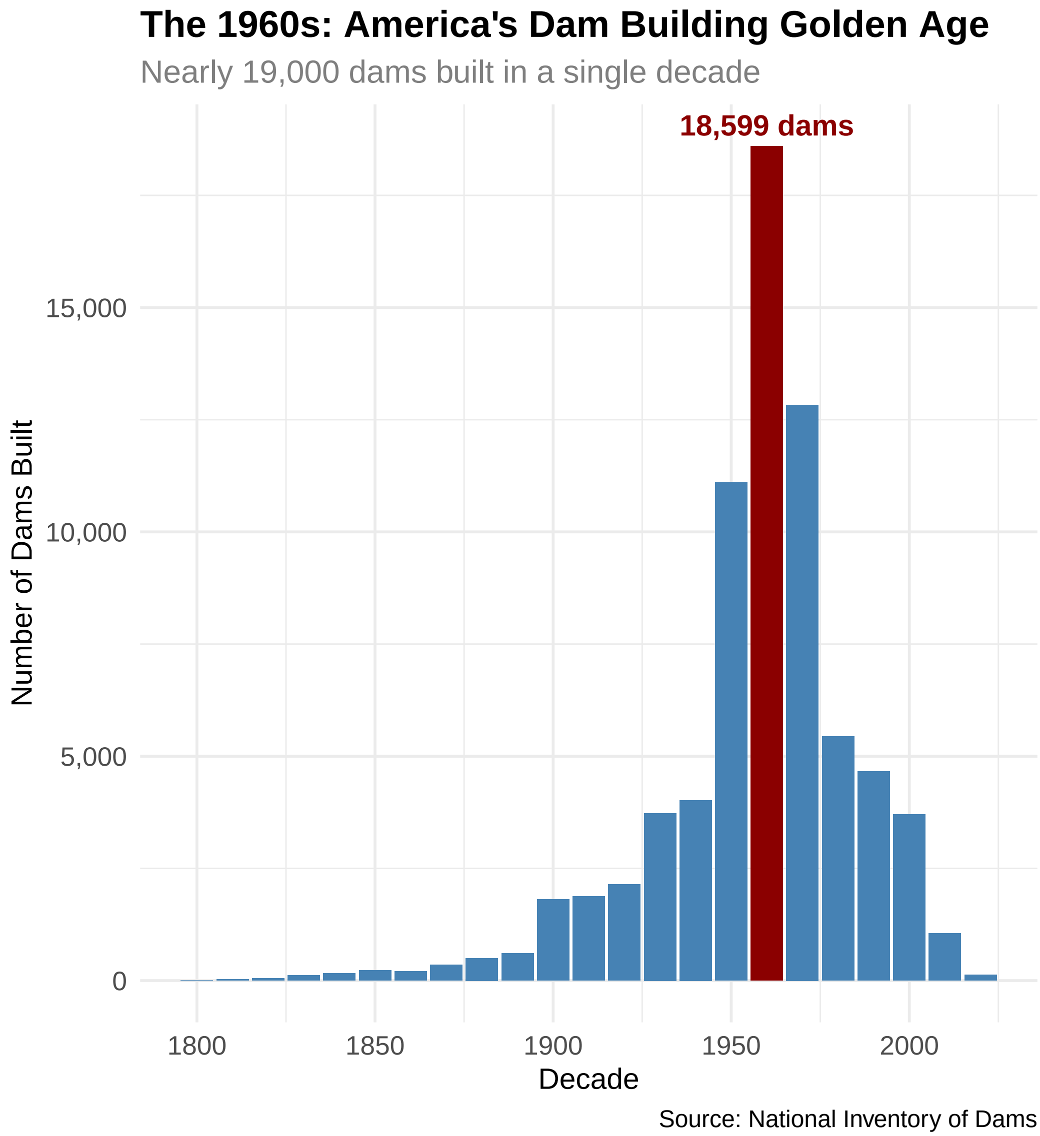

Figure 1: Dam Construction by Decade: The 1960s Boom

Fun Fact: More dams were built in the **1960s alone** than in the previous 160 years combined!

What Are Dams Actually Used For?

Contrary to popular belief, **flood control** isn't the primary purpose of most American dams. Recreation is the most common purpose.

```
purpose_data <- dat |>
  count(primary_purpose) |>
  arrange(desc(n)) |>
  slice(1:10) |>
  mutate(
    primary_purpose = case_when(
      primary_purpose == "Fire Protection, Stock, Or Small Fish Pond" ~ "Fire/Stock/Fish Pond
      TRUE ~ primary_purpose
    ),
    primary_purpose = reorder(primary_purpose, n),
    percentage = n / sum(n) * 100
  )

ggplot(purpose_data, aes(x = primary_purpose, y = n)) +
  geom_col(fill = "forestgreen", alpha = 0.8) +
  geom_text(aes(label = paste0(format(n, big.mark = ","), "\n(", round(percentage, 1), "%)"))
            hjust = -0.1, size = 3.5) +
  coord_flip() +
  scale_y_continuous(labels = comma_format(), expand = expansion(mult = c(0, 0.15))) +
  labs(
    title = "Recreation Dominates American Dam Purposes",
    subtitle = "Top 10 primary purposes for the nation\'s 92,428 dams",
    x = "Primary Purpose",
    y = "Number of Dams",
    caption = "Source: National Inventory of Dams"
  ) +
  theme_dam
```

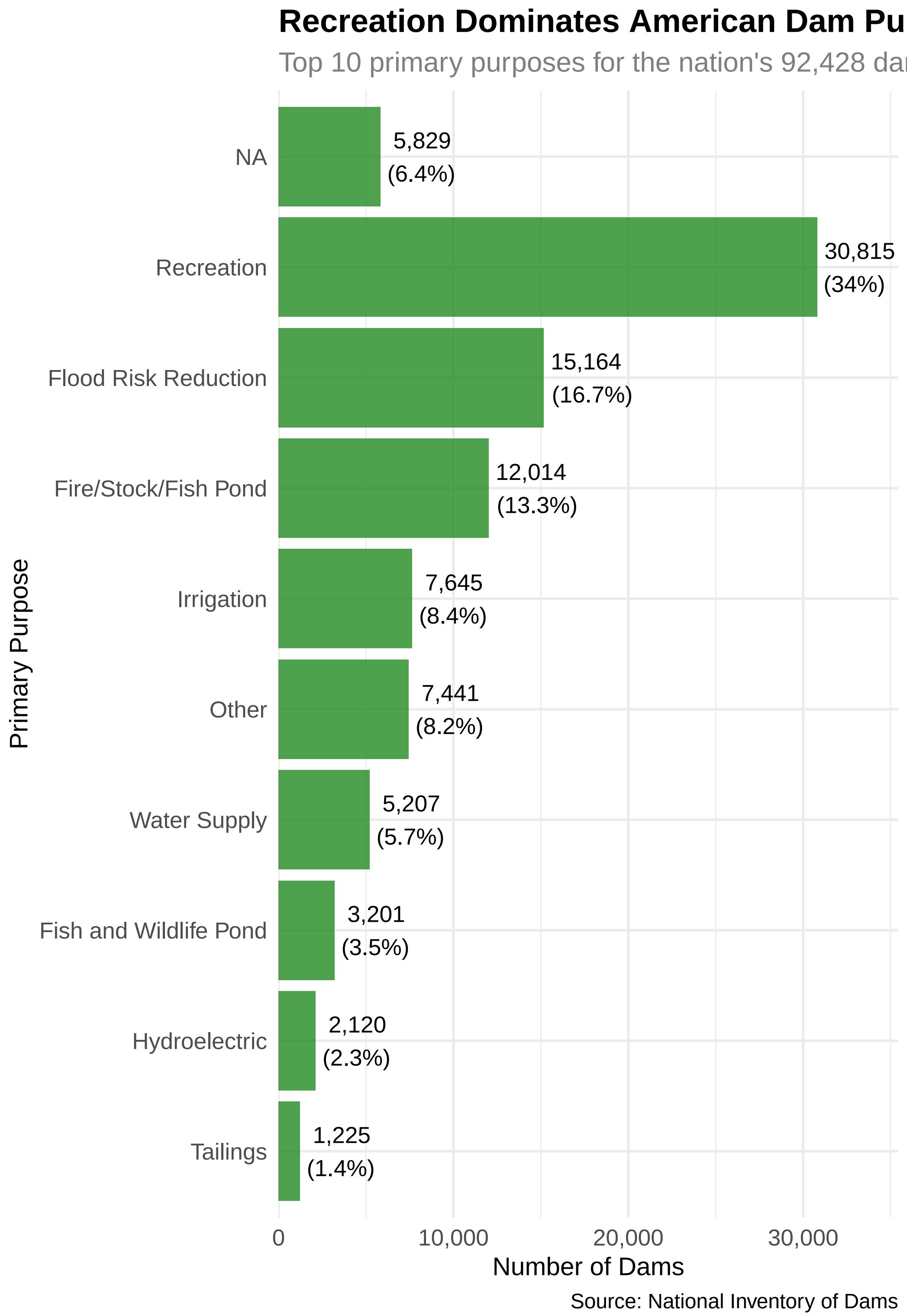

Figure 2: Primary Purposes of American Dams

The Surprising Recreation Story

30,815 dams (**33.3%**) are primarily used for recreation. This includes:

- Swimming and boating
- Fishing ponds

- Campground water features
- Golf course ponds

Safety First: High-Hazard Dams

The safety implications are sobering: **16,843** dams are classified as **"High Hazard"**, meaning their failure would likely cause loss of life.

```
hazard_data <- dat |>
  filter(hazard_potential_classification == "High") |>
  count(state) |>
  arrange(desc(n)) |>
  slice(1:15) |>
  mutate(
    state = reorder(state, n),
    danger_level = case_when(
      n >= 1500 ~ "Extreme Risk",
      n >= 800 ~ "High Risk",
      TRUE ~ "Moderate Risk"
    )
  )

ggplot(hazard_data, aes(x = state, y = n, fill = danger_level)) +
  geom_col() +
  geom_text(aes(label = format(n, big.mark = ",")), hjust = -0.1, size = 3) +
  coord_flip() +
  scale_fill_manual(
    values = c("Extreme Risk" = "darkred", "High Risk" = "orange", "Moderate Risk" = "gold")
    name = "Risk Level"
  ) +
  scale_y_continuous(labels = comma_format(), expand = expansion(mult = c(0, 0.12))) +
  labs(
    title = "States with the Most High-Hazard Dams",
```

```
  subtitle = "These dams could cause loss of life if they fail",
  x = "State",
  y = "Number of High-Hazard Dams",
  caption = "Source: National Inventory of Dams"
) +
theme_dam
```

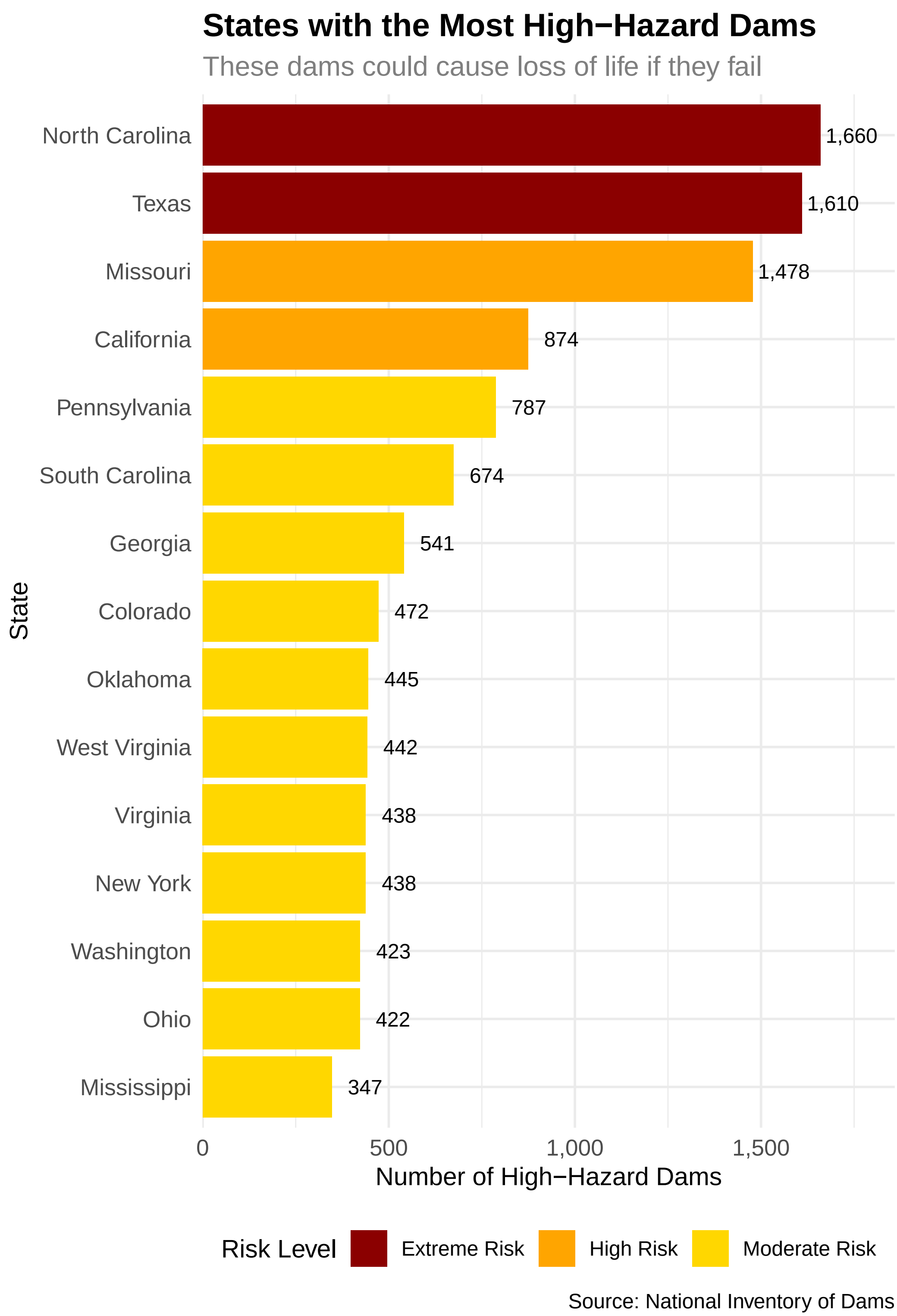

Figure 3: High-Hazard Dams by State

⚠ Safety Spotlight

North Carolina leads with **1,660** high-hazard dams, followed by **Texas** with **1,610**.

Tallest Dams: Engineering Marvels

America's Skyscrapers of Water

```
tallest_data <- dat |>
  select(dam_name, state, dam_height_ft) |>
  filter(!is.na(dam_height_ft)) |>
  arrange(desc(dam_height_ft)) |>
  slice(1:10) |>
  mutate(
    dam_name = reorder(dam_name, dam_height_ft),
    dam_type = case_when(
      grepl("Slurry|Refuse|Tailings", dam_name, ignore.case = TRUE) ~ "Mining Waste",
      dam_name %in% c("Hoover Dam", "Glen Canyon Dam", "Oroville") ~ "Major Federal",
      TRUE ~ "Other"
    )
  )

ggplot(tallest_data, aes(x = dam_name, y = dam_height_ft, fill = dam_type)) +
  geom_col() +
  geom_text(aes(label = paste0(dam_height_ft, " ft")), hjust = -0.1, size = 3) +
  coord_flip() +
  scale_fill_manual(
    values = c("Major Federal" = "navy", "Mining Waste" = "brown", "Other" = "gray60"),
    name = "Dam Type"
  ) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.12))) +
  labs(
    title = "America\'s Tallest Dams: A Mix of Federal Projects and Mining Waste",
    subtitle = "Several mining waste dams rank among the nation\'s tallest structures",
    x = "Dam Name",
    y = "Height (Feet)",
    caption = "Source: National Inventory of Dams"
```

```
) +
  theme_dam
```

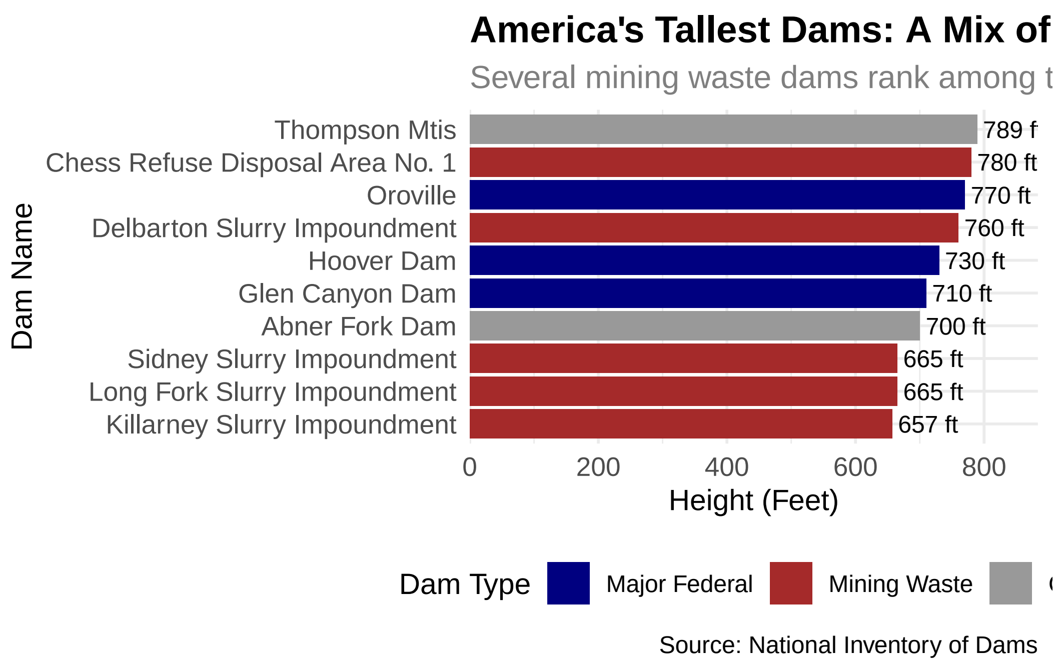

Figure 4: The 10 Tallest Dams in America

Surprising finding: Several of America's tallest "dams" are actually **mining waste containment structures** in Kentucky and West Virginia!

Who Owns America's Dams?

The ownership story might surprise you: **65%** of all dams are **privately owned**.

```
ownership_data <- dat |>
  count(primary_owner_type) |>
  filter(!is.na(primary_owner_type), primary_owner_type != "Not Listed") |>
  arrange(desc(n)) |>
  mutate(
    percentage = n / sum(n) * 100,
```

```
    label = paste0(primary_owner_type, "\\n", format(n, big.mark = ","), " dams\\n(", round(p
  )

ggplot(ownership_data, aes(x = "", y = n, fill = primary_owner_type)) +
  geom_col(width = 1) +
  coord_polar("y", start = 0) +
  scale_fill_brewer(type = "qual", palette = "Set3") +
  theme_void() +
  theme(legend.position = "right") +
  labs(
    title = "Private Ownership Dominates American Dams",
    subtitle = "Nearly two-thirds of dams are privately owned",
    fill = "Owner Type",
    caption = "Source: National Inventory of Dams"
  )
```

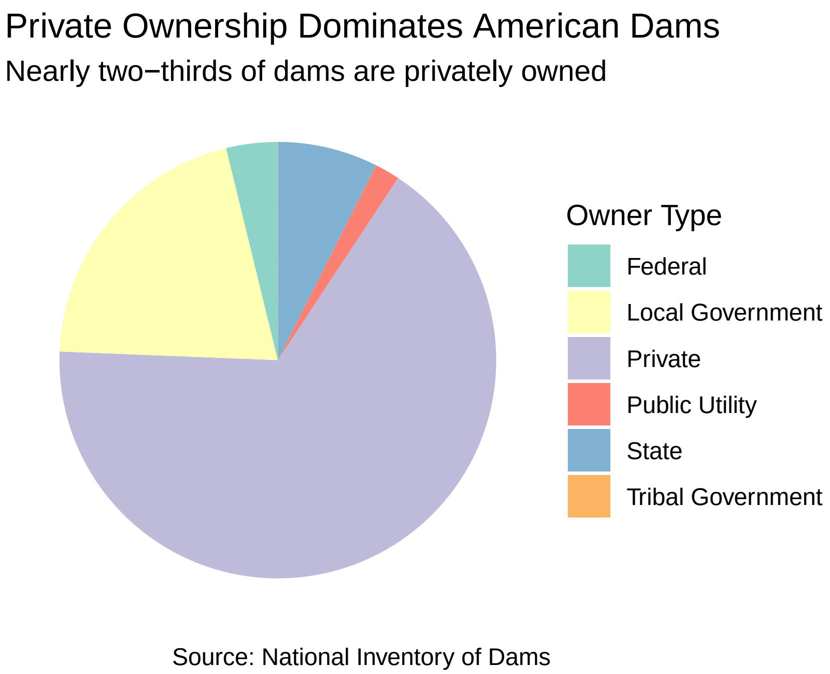

Figure 5: Dam Ownership Distribution

Hydroelectric Hotspots

Where does America generate hydroelectric power? The answer might surprise you - it's not just the Pacific Northwest!

```r
hydro_data <- dat |>
  filter(primary_purpose == "Hydroelectric") |>
  count(state) |>
  arrange(desc(n)) |>
  slice(1:12) |>
  mutate(
    state = reorder(state, n),
    region = case_when(
      state %in% c("Maine", "New Hampshire", "Vermont", "Massachusetts") ~ "New England",
      state %in% c("New York") ~ "Mid-Atlantic",
      state %in% c("California", "Washington", "Oregon") ~ "West Coast",
      TRUE ~ "Other"
    )
  )

ggplot(hydro_data, aes(x = state, y = n, fill = region)) +
  geom_col() +
  geom_text(aes(label = n), hjust = -0.1, size = 3) +
  coord_flip() +
  scale_fill_manual(
    values = c("New England" = "darkgreen", "Mid-Atlantic" = "navy",
               "West Coast" = "orange", "Other" = "gray60"),
    name = "Region"
  ) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.12))) +
  labs(
    title = "New York Leads in Hydroelectric Dams",
    subtitle = "New England states dominate small-scale hydroelectric generation",
    x = "State",
    y = "Number of Hydroelectric Dams",
    caption = "Source: National Inventory of Dams"
  ) +
  theme_dam
```

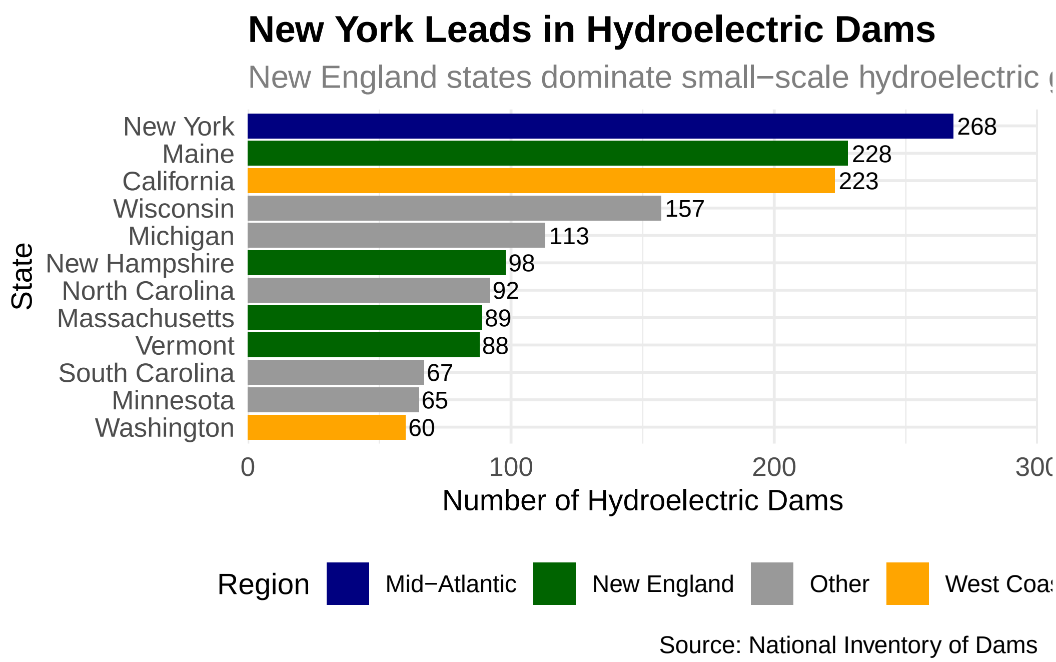

Figure 6: Hydroelectric Dams by State

Key insight: **New England states** (Maine, New Hampshire, Vermont, Massachusetts) collectively have more hydroelectric dams than the traditionally hydro-focused Western states!

Average Dam Heights by State

Which states build the tallest dams on average?

```
height_data <- dat |>
  filter(!is.na(dam_height_ft), dam_height_ft > 0) |>
  group_by(state) |>
  summarise(
    avg_height = round(mean(dam_height_ft), 1),
    count = n(),
    .groups = "drop"
  ) |>
  filter(count >= 50) |>
  arrange(desc(avg_height)) |>
  slice(1:15) |>
```

```r
  mutate(
    state = reorder(state, avg_height),
    highlight = state %in% c("Washington", "West Virginia", "Colorado")
  )

ggplot(height_data, aes(x = state, y = avg_height, fill = highlight)) +
  geom_col(show.legend = FALSE) +
  geom_text(aes(label = paste0(avg_height, " ft")), hjust = -0.1, size = 3) +
  coord_flip() +
  scale_fill_manual(values = c("FALSE" = "steelblue", "TRUE" = "darkred")) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.12))) +
  labs(
    title = "Washington State Builds the Tallest Dams on Average",
    subtitle = "Average height for states with 50+ dams in the database",
    x = "State",
    y = "Average Dam Height (Feet)",
    caption = "Source: National Inventory of Dams"
  ) +
  theme_dam
```

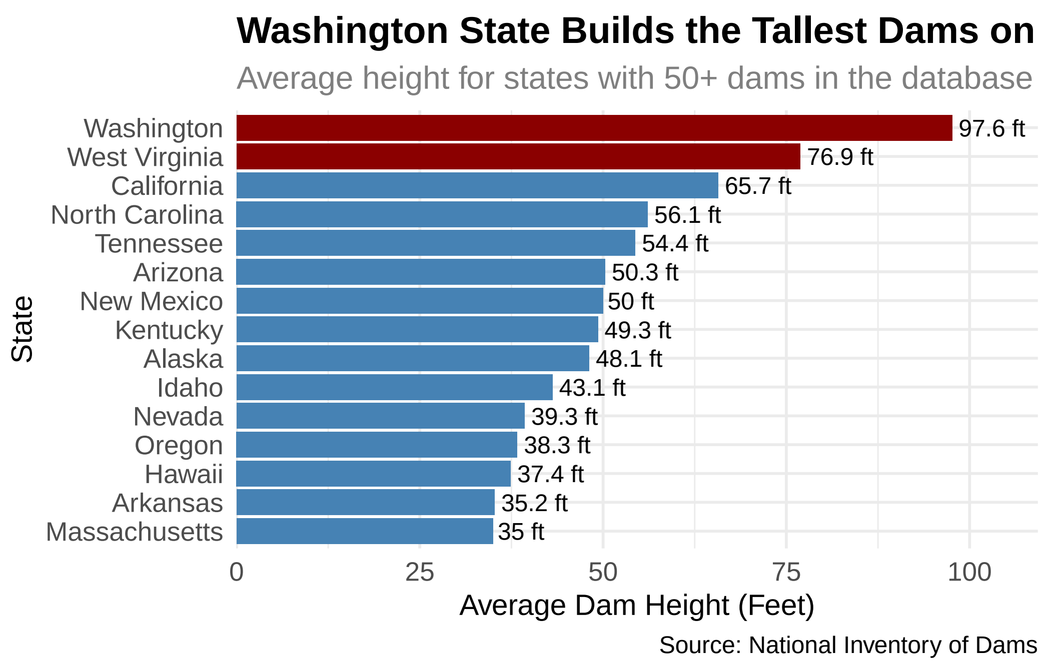

Figure 7: Average Dam Height by State (States with 50+ dams)

Strange and Fascinating Facts

The Mystery of “Year 0” Dams

There are **4** dams in the dataset with a completion year of **0**. Let's investigate:

```
year_zero_dams <- dat |>
  filter(year_completed == 0) |>
  select(dam_name, state, primary_purpose) |>
  slice(1:10)

if(nrow(year_zero_dams) > 0) {
  kable(year_zero_dams, caption = "Sample of Mysterious \'Year 0\' Dams")
} else {
  cat("No dams with year 0 found in this dataset.")
}
```

Table 1: Sample of Mysterious ‘Year 0’ Dams

dam_name	state	primary_purpose
Wailuku Water Reservoir 10	Hawaii	Irrigation
Wailuku Water Reservoir 6	Hawaii	Irrigation
Halaula Reservoir	Hawaii	Irrigation
Pinau Reservoir	Hawaii	Other

Florida's Mining Surprise

Let's look at states with the most **tailings** (mining waste) dams:

```
tailings_data <- dat |>
  filter(primary_purpose == "Tailings") |>
  count(state) |>
  arrange(desc(n)) |>
  slice(1:10) |>
  mutate(state = reorder(state, n))

ggplot(tailings_data, aes(x = state, y = n)) +
  geom_col(fill = "brown", alpha = 0.8) +
```

```
geom_text(aes(label = n), hjust = -0.1, size = 3) +
coord_flip() +
scale_y_continuous(expand = expansion(mult = c(0, 0.12))) +
labs(
  title = "Mining Waste Dams by State",
  subtitle = "Tailings dams for containing mining waste",
  x = "State",
  y = "Number of Tailings Dams",
  caption = "Source: National Inventory of Dams"
) +
theme_dam
```

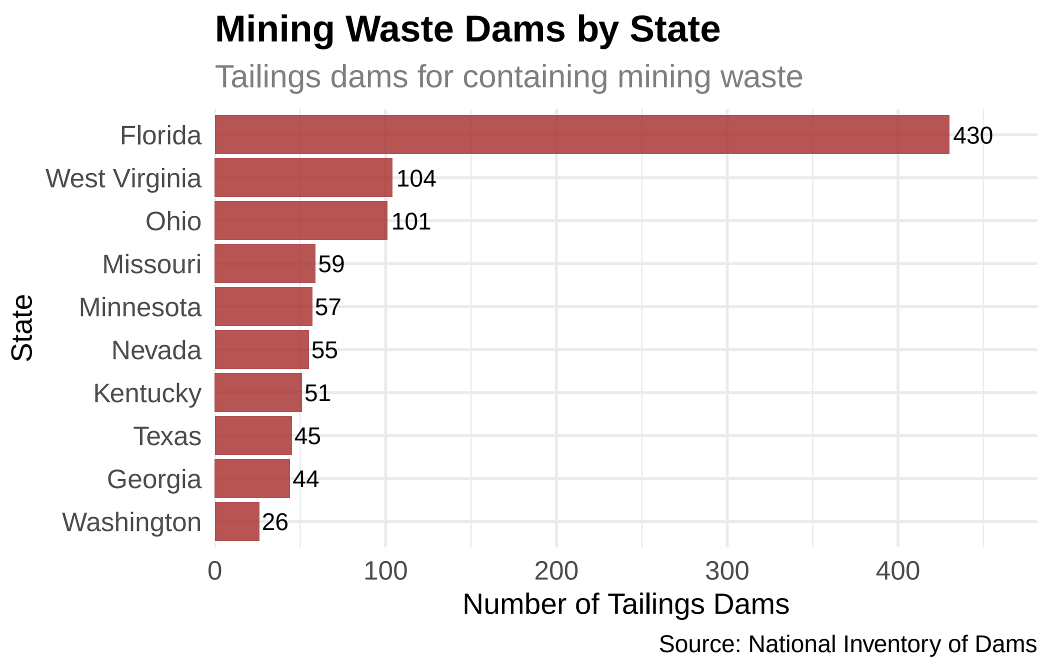

Figure 8: States with Most Mining Waste Dams

Key Takeaways

> **What We Learned**
>
> 1. **Infrastructure Boom**: The 1960s were the golden age of dam construction
> 2. **Recreation Rules**: Most dams serve recreational purposes, not flood control
> 3. **Safety Concerns**: Nearly 17,000 high-hazard dams need monitoring
> 4. **Private Property**: Nearly 2/3 of dams are privately owned
> 5. **Hydro Surprise**: New England, not the West, leads in hydroelectric dam count
> 6. **Height Champions**: Washington state builds the tallest dams on average
> 7. **Mining Impact**: Mining waste dams are concentrated in specific states

---

Dataset Summary

Data Source: National Inventory of Dams (NID)
Maintained by: U.S. Army Corps of Engineers
Total Records: 92,428
Variables: 83
Coverage: All 50 US states plus territories
Analysis Date: 2025-06-20

Data and Analysis

This analysis uses data from the **National Inventory of Dams (NID)**, maintained by the U.S. Army Corps of Engineers. The NID is a comprehensive database containing information on dams throughout the United States and its territories.

This analysis was prepared with the assistance of Anthropic Claude 4 Sonnet.

> **About the National Inventory of Dams**
>
> The NID was established following the National Dam Safety Act of 1972 and serves as a key resource for dam safety, emergency preparedness, and water resource management. The database is regularly updated with information submitted by state dam safety agencies and federal agencies.