MC_SMCVAS Baseline Analysis

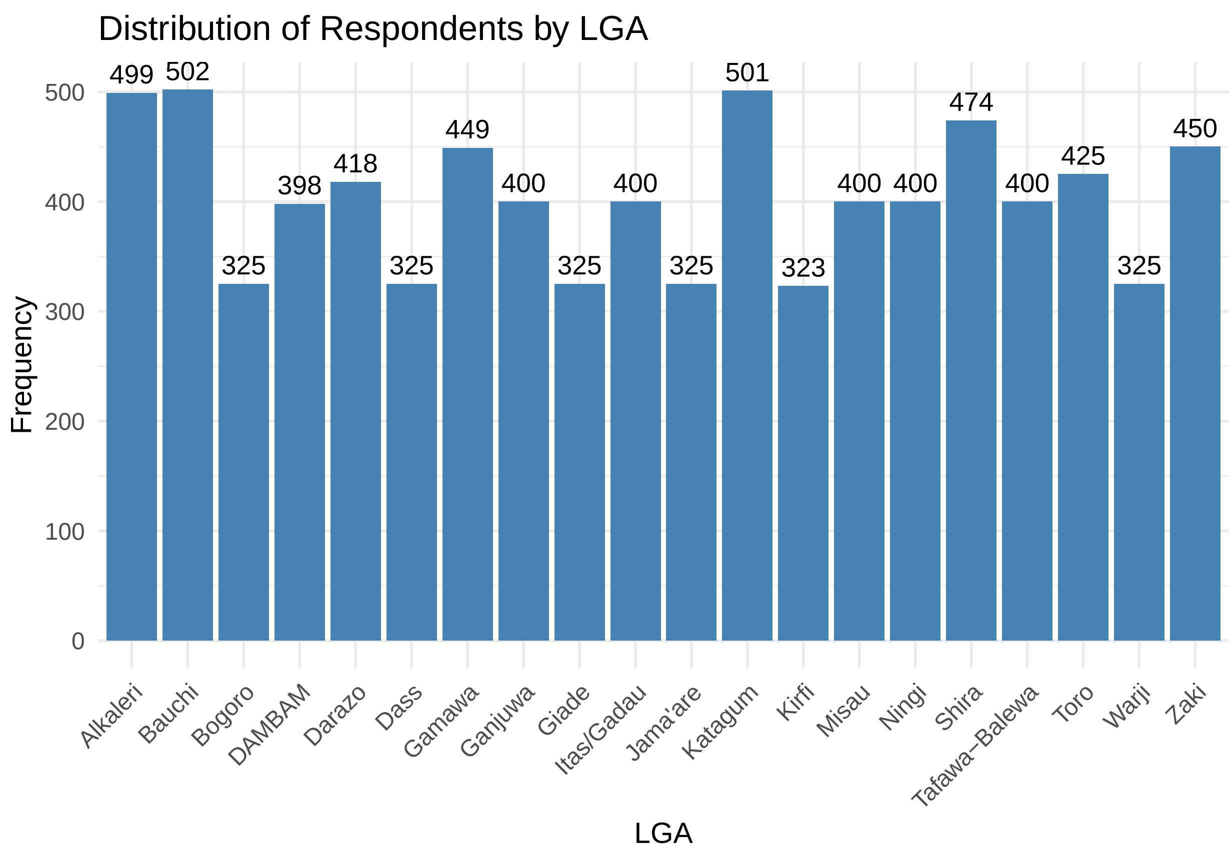
AbdulHafiz Abba

2025-06-17

```r
library(readxl)

df <- read_excel("Dataset/SMC_VAS BAUCHI BASELINE SURVEY 2025_Play.xlsx")
library(dplyr)
library(tidyr)
library(ggplot2)
library(scales)
```

```
##             LGA Frequency Percent
##        Alkaleri       499     6.2
##          Bauchi       502     6.2
##          Bogoro       325     4.0
##          DAMBAM       398     4.9
##          Darazo       418     5.2
##            Dass       325     4.0
##          Gamawa       449     5.6
##         Ganjuwa       400     5.0
##           Giade       325     4.0
##      Itas/Gadau       400     5.0
##        Jama'are       325     4.0
##         Katagum       501     6.2
##           Kirfi       323     4.0
##           Misau       400     5.0
##           Ningi       400     5.0
##           Shira       474     5.9
##   Tafawa-Balewa       400     5.0
##            Toro       425     5.3
##           Warji       325     4.0
##            Zaki       450     5.6
##           Total      8064   100.0
```



Analyze Relationships of the Wealth Index With Other Key Variables

Creating the Coverage Variables

```
df <- df %>%
  mutate(
    VAS_coverage = ifelse(`Did [selected_chi_name] received vitamin A dose within the last 6 months from
    Deworming_coverage = ifelse(`Did [selected_chi_name] receive a deworming tablet (during the last MN
    MUAC_coverage = ifelse(`Did [selected_chi_name] received a MUAC screening (during the last MNCHW ca
    Immunization_coverage = ifelse(`Question: Did [selected_chi_name] receive any vaccine during the la

  )
```

```
coverage_services <- df %>%
  group_by(LGA) %>%
  summarize(
    VAS = mean(VAS_coverage, na.rm = TRUE),
    Deworming = mean(Deworming_coverage, na.rm = TRUE),
    MUAC = mean(MUAC_coverage, na.rm = TRUE),
    Immunization = mean(Immunization_coverage, na.rm = TRUE),
    n = n()
  )

print(coverage_services)
```

```
## # A tibble: 20 x 6
##    LGA            VAS Deworming   MUAC Immunization     n
##    <chr>        <dbl>     <dbl>  <dbl>        <dbl> <int>
##  1 Alkaleri     0.423     0.357 0.265        0.971    499
##  2 Bauchi       0.510     0.414 0.416        0.558    502
##  3 Bogoro       0.452     0.418 0.222        0.753    325
##  4 DAMBAM       0.374     0.317 0.221        0.917    398
##  5 Darazo       0.667     0.569 0.514        0.545    418
##  6 Dass         0.566     0.366 0.277        0.737    325
##  7 Gamawa       0.122     0.236 0.0512       0.652    449
##  8 Ganjuwa      0.415     0.398 0.152        0.378    400
##  9 Giade        0.535     0.357 0.228        0.417    325
## 10 Itas/Gadau   0.462     0.372 0.0975       0.286    400
## 11 Jama'are     0.274     0.172 0.129        0.0909   325
## 12 Katagum      0.184     0.114 0.0699       1        501
## 13 Kirfi        0.297     0.393 0.307        0.8      323
## 14 Misau        0.472     0.412 0.308        0.846    400
## 15 Ningi        0.37      0.268 0.14         0.914    400
## 16 Shira        0.228     0.213 0.148        0.821    474
## 17 Tafawa-Balewa 0.392    0.315 0.252        0.818    400
## 18 Toro         0.536     0.381 0.334        0.444    425
## 19 Warji        0.36      0.317 0.249        0.889    325
## 20 Zaki         0.291     0.213 0.173        0.448    450
```

Coverage of Key Child Health Interventions Across LGAs

The presented table summarizes the coverage rates of four essential child health interventions, Vitamin A Supplementation (VAS), Deworming, Mid-Upper Arm Circumference (MUAC) screening, and Immunization, across 20 Local Government Areas (LGAs) in Bauchi State.

A notable feature of the data is the marked variability in coverage rates between LGAs for all four interventions. Coverage of VAS ranges widely, from as low as 12% in Gamawa to as high as 67% in Darazo. Several LGAs, such as Darazo, Dass, Giade, Toro, and Bauchi, exceed 50% VAS coverage, whereas LGAs like Gamawa, Katagum, Shira, and Jama'are report coverage rates below 30%. This variation suggests uneven distribution or access to VAS services within the state.

Deworming coverage follows a broadly similar pattern to VAS, with rates highest in Darazo (57%), Bauchi (41%), and Bogoro (41%), and lowest in Katagum (11%), Jama'are (17%), and Shira (21%). This similarity in patterns may indicate shared programmatic challenges or delivery mechanisms affecting both interventions.

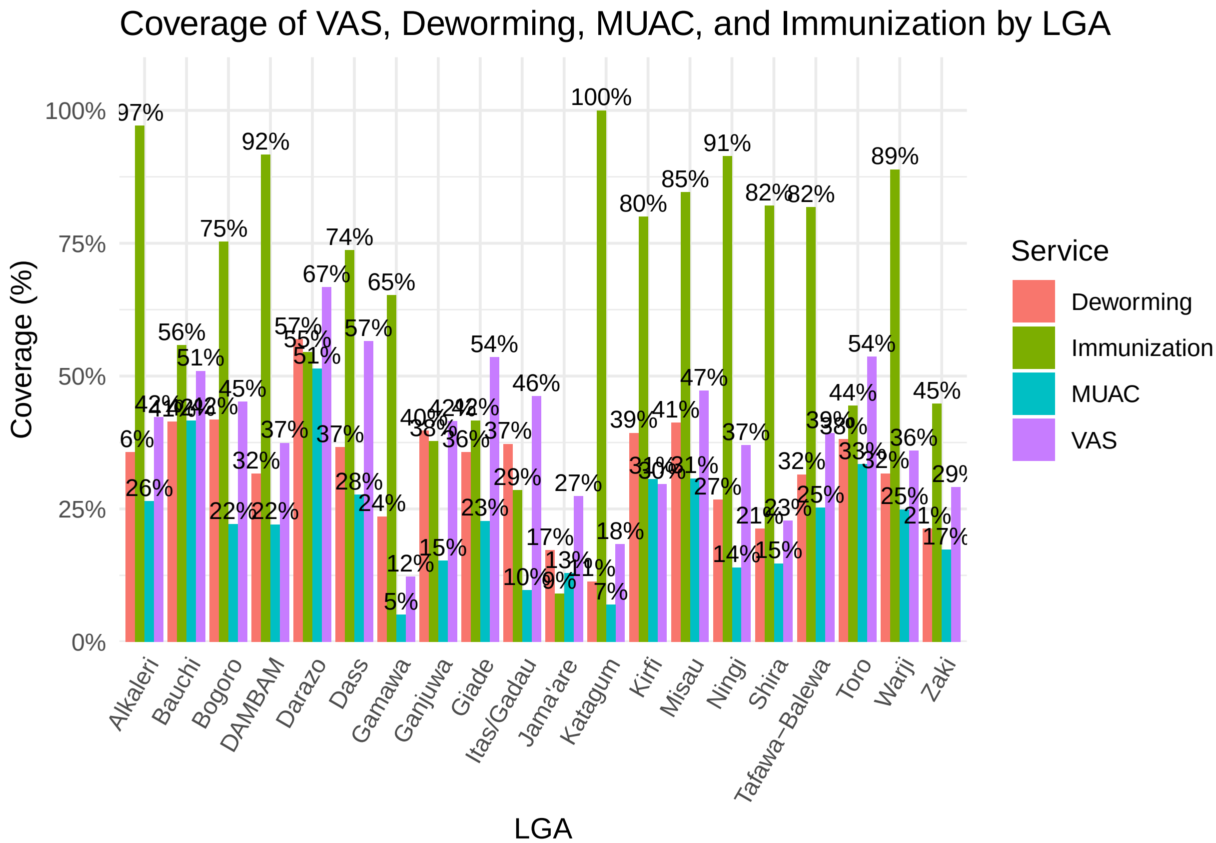
MUAC screening coverage is consistently the lowest among the four interventions across most LGAs. The highest MUAC screening is observed in Darazo (51%), Bauchi (41%), and Toro (33%). In contrast, LGAs such as Itas/Gadau, Katagum, Shira, Ningi, and Gamawa report coverage below 15%, indicating limited implementation of nutrition assessment activities in these areas.

Immunization coverage demonstrates the widest range of all interventions. Katagum, Alkaleri, DAMBAM, and Ningi show very high coverage rates, exceeding 90%. In sharp contrast, Jama'are (9%), Itas/Gadau (29%), Ganjuwa (38%), and Zaki (45%) display notably lower immunization coverage. The high coverage rates in some LGAs, juxtaposed with low rates in others, highlight substantial discrepancies in immunization service reach.

Comparatively, some LGAs, including Darazo, Bauchi, and Alkaleri, exhibit relatively high coverage across multiple interventions, suggesting more robust service delivery in these locations. Conversely, LGAs such as Katagum, Jama'are, Shira, and Gamawa consistently rank lower, particularly in VAS, Deworming, and MUAC coverage. Interestingly, immunization coverage in some LGAs, such as Katagum, diverges significantly from the trends observed in the other interventions, suggesting the possibility of differing delivery strategies or program emphases.

```r
# Table of Coverage by LGA and Service for Plotting
coverage_long <- coverage_services %>%
  pivot_longer(
    cols = c(VAS, Deworming, MUAC, Immunization),
    names_to = "Service",
    values_to = "Coverage"
  )

# Visualize Coverage by LGA and Service
ggplot(coverage_long, aes(x = LGA, y = Coverage, fill = Service)) +
  geom_col(position = position_dodge(width = 0.9)) +
  geom_text(
    aes(label = scales::percent(Coverage, accuracy = 1)),
    position = position_dodge(width = 0.9),
    vjust = -0.3,
    size = 3
  ) +
  scale_y_continuous(labels = percent_format(accuracy = 1), expand = expansion(mult = c(0, 0.1))) +
  labs(
    title = "Coverage of VAS, Deworming, MUAC, and Immunization by LGA",
    x = "LGA",
    y = "Coverage (%)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, hjust = 1))
```



Statistical Test for Difference in Coverage Across LGAs

Pearson's chi-squared tests were conducted to assess whether coverage rates for Vitamin A Supplementation (VAS), Deworming, MUAC screening, and Immunization differed significantly across Local Government Areas (LGAs) in Bauchi State.

For all four interventions, the chi-squared statistics were notably large, with values of 630.91 for VAS (df = 19), 402.38 for Deworming (df = 19), 624.52 for MUAC screening (df = 19), and 138.08 for Immunization (df = 19). In each case, the associated p-value was less than 2.2e-16.

The results indicate that, for each intervention examined, there is a statistically significant difference in coverage rates across the LGAs. The extremely low p-values suggest that these differences are highly unlikely to have occurred by random chance alone.

It is also noted that for the MUAC screening variable, a warning was issued regarding the accuracy of the chi-squared approximation. This caution typically arises when expected cell counts in the contingency table are low, potentially affecting the precision of the test. Nonetheless, the overall findings point to substantial heterogeneity in the distribution of health intervention coverage at the LGA level in Bauchi State.

```r
# VAS
chisq_vas <- chisq.test(table(df$LGA, df$VAS_coverage))
print(chisq_vas)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(df$LGA, df$VAS_coverage)
## X-squared = 630.91, df = 19, p-value < 2.2e-16
```

```r
# Deworming
chisq_deworming <- chisq.test(table(df$LGA, df$Deworming_coverage))
print(chisq_deworming)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(df$LGA, df$Deworming_coverage)
## X-squared = 402.38, df = 19, p-value < 2.2e-16
```

```r
# MUAC
chisq_muac <- chisq.test(table(df$LGA, df$MUAC_coverage))
print(chisq_muac)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(df$LGA, df$MUAC_coverage)
## X-squared = 624.52, df = 19, p-value < 2.2e-16
```

```r
# Immunization
chisq_immun <- chisq.test(table(df$LGA, df$Immunization_coverage))
```

```
## Warning in chisq.test(table(df$LGA, df$Immunization_coverage)): Chi-squared
## approximation may be incorrect
```

```r
print(chisq_immun)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(df$LGA, df$Immunization_coverage)
## X-squared = 138.08, df = 19, p-value < 2.2e-16
```

Wealth Index Analysis

```r
df <- df %>%
  mutate(
    water_source_score = case_when(
      HH_watersource == "Piped water" ~ 3,
      HH_watersource == "Bottled water" ~ 5,
      HH_watersource == "Sachet water/pure water" ~ 4,
      HH_watersource == "Rainwater" ~ 0,
      HH_watersource == "Tanker truck" ~ 1,
      HH_watersource == "Cart with small tank" ~ 0,
      HH_watersource == "Dug well" ~ 0,
      HH_watersource == "Surface water (river/dam/lake/pond/stream/canal/irrigation channel)" ~ 0,
      HH_watersource == "Others" ~ 0,
      HH_watersource == "" | is.na(HH_watersource) ~ 0,
```

```r
      HH_watersource == "Water from spring" ~ 0,
      TRUE ~ 0
    ),
    toilet_type_score = case_when(
      HH_Toilet_type == "FLUSH OR POUR FLUSH TOILET" ~ 3,
      HH_Toilet_type == "PIT LATRINE" ~ 1,
      HH_Toilet_type == "" | is.na(HH_Toilet_type) ~ 0,
      TRUE ~ 0
    ),
    electricity_score = ifelse(HH_appliance_Electricity == "Yes", 1, 0),
    refrigerator_score = ifelse(HH_appliance_Refregerator == "Yes", 1, 0),
    radio_score = ifelse(HH_appliance_Radio == "Yes", 1, 0),
    TV_score = ifelse(HH_appliance_Television == "Yes", 1, 0),
    cable_TV_score = ifelse(HH_appliance_Cable_TV == "Yes", 1, 0),
    AC_score = ifelse(HH_appliance_Air_Conditioner == "Yes", 1, 0),
    computer_score = ifelse(HH_appliance_Computer == "Yes", 1, 0),
    fan_score = ifelse(HH_appliance_Fan == "Yes", 1, 0),
    non_appliance_score = ifelse(HH_appliance_None == "Yes", 1, 0),
    iron_score = ifelse(`HH_appliance_Electronic Iron` == "Yes", 1, 0),
    generating_set_score = ifelse(`HH_appliance_Generating Set` == "Yes", 1, 0),
    mobile_Phone_score = ifelse(`HH_appliance_Mobile telephone` == "Yes", 1, 0),
    non_mobile_phone_score = ifelse(`HH_appliance_Non_Mobile telephone` == "Yes", 1, 0),
    floor_type_score = case_when(
      HH_floor_type == "FINISHED FLOOR" ~ 3,
      HH_floor_type == "NATURAL FLOOR" ~ 0,
      TRUE ~ 0
    ),
    fuel_source_score = case_when(
      HH_fuel_source == "ELECTRICITY" ~ 5,
      HH_fuel_source == "AGRICULTURAL CROP" ~ 0,
      HH_fuel_source == "ANIMAL DUNG" ~ 0,
      HH_fuel_source == "BIOGAS" ~ 3,
      HH_fuel_source == "CHARCOAL" ~ 2,
      HH_fuel_source == "COAL, LIGNITE" ~ 2,
      HH_fuel_source == "LIQUID PROPANE GAS/CYLINDER" ~ 4,
      HH_fuel_source == "NATURAL GAS" ~ 4,
      HH_fuel_source == "STRAW/SHRUBS/GRASS/SAWDUST" ~ 1,
      HH_fuel_source == "WOOD" ~ 1,
      HH_fuel_source == "" | is.na(HH_fuel_source) ~ 0,
      TRUE ~ 0
    )
  )
```

```r
df <- df %>%
  mutate(
    wealth_score = rowSums(
      across(
        c(
          water_source_score,
          toilet_type_score,
          electricity_score,
          refrigerator_score,
          radio_score,
```

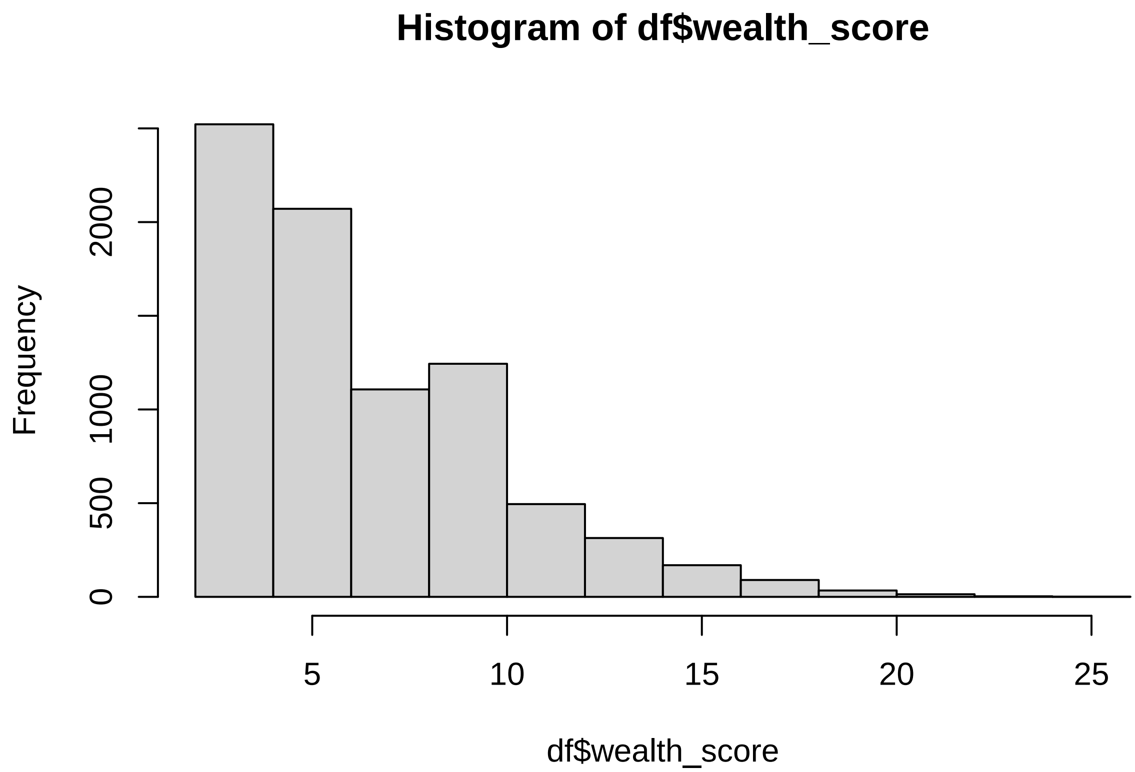
```r
          TV_score,
          cable_TV_score,
          AC_score,
          computer_score,
          fan_score,
          non_appliance_score,
          iron_score,
          generating_set_score,
          mobile_Phone_score,
          non_mobile_phone_score,
          floor_type_score,
          fuel_source_score
        )
      ),
      na.rm = TRUE
    )
  )
```

Summarize the Wealth Score Distribution

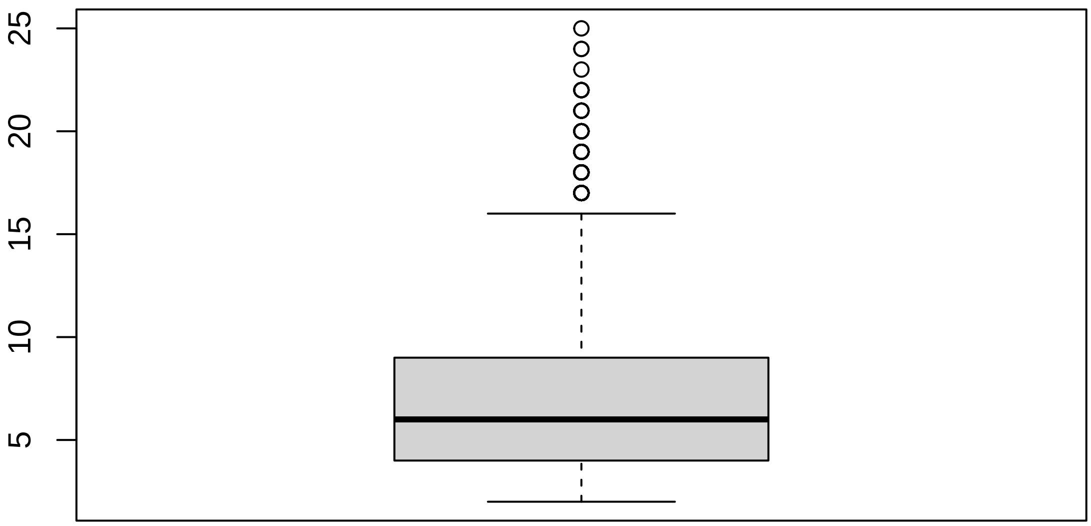
```r
summary(df$wealth_score)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00    4.00    6.00    6.86    9.00   25.00
```

```r
hist(df$wealth_score)
```




```
boxplot(df$wealth_score)
```




```r
df <- df %>%
  mutate(
    wealth_quintile = ntile(wealth_score, 5),
    wealth_quintile_label = factor(
      wealth_quintile,
      levels = 1:5,
      labels = c("Poorest", "Poor", "Middle", "Rich", "Richest")
    )
  )
```

Compare Coverage by Wealth Quintile

```r
library(janitor)
```

```
## Warning: package 'janitor' was built under R version 4.4.3
```

```
##
## Attaching package: 'janitor'
```

```
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```r
df %>%
  tabyl(wealth_quintile_label, VAS_coverage) %>%
  adorn_percentages("row") %>%
  adorn_pct_formatting()
```

```
##  wealth_quintile_label     0     1
##                Poorest 68.2% 31.8%
##                   Poor 58.8% 41.2%
##                 Middle 65.5% 34.5%
##                   Rich 57.8% 42.2%
##                Richest 53.6% 46.4%
```

Barplot of VAS coverage by wealth:

```r
library(ggplot2)
df %>%
  group_by(wealth_quintile_label) %>%
  summarize(VAS_coverage_rate = mean(VAS_coverage, na.rm = TRUE)) %>%
  ggplot(aes(x = wealth_quintile_label, y = VAS_coverage_rate, fill = wealth_quintile_label)) +
  geom_col() +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "VAS Coverage by Wealth Quintile", x = "Wealth Quintile", y = "Coverage (%)") +
  theme_minimal() +
  theme(legend.position = "none")
```