# Summary of Descriptive Statistics and Interpretations of Predictors for Adverse Pregnancy Outcomes (APO)

**Dataset**: 2022 Kenya Demographic and Health Survey (DHS)  
**Outcome Variable**: apo\_general (NonAPO = No Adverse Pregnancy Outcome, APO = Adverse Pregnancy Outcome)

This document summarizes the descriptive statistics and interpretations of predictors analyzed for their association with adverse pregnancy outcomes (APO) in the 2022 Kenya DHS dataset. Each predictor is presented with its descriptive statistics, comparison between NonAPO and APO groups, and interpretation in the context of APO risk and the 2022 Kenya DHS findings.

## 1. Average Timing of First Antenatal Check (avg\_timing\_first\_antenatal\_check)

* **Description**: Average month of the first antenatal care visit across all pregnancies (in months, capped at 9 months; values > 9 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Mean = 4.56 months.
  + APO: Mean = 4.21 months.
  + Difference: -0.35 months (APO women start antenatal care earlier).
* **Interpretation**:
  + APO women initiate antenatal care earlier (4.21 months vs. 4.56 months), which is unexpected since earlier antenatal care is typically associated with better outcomes. This might reflect that APO women, who are more urban, wealthier, and educated, have better access to care and start earlier. However, earlier visits might also indicate higher-risk pregnancies (e.g., history of APO prompting earlier care).
  + **Context with 2022 Kenya DHS**: The DHS reports that 66% of women had their first antenatal visit in the first trimester (before 4 months), but our sample shows later averages (4.21–4.56 months), possibly because our data includes all pregnancies, not just the most recent, and averages across multiple births may include later visits.

## 2. Average Number of Antenatal Visits (avg\_num\_antenatal\_visits)

* **Description**: Average number of antenatal care visits across all pregnancies (capped at 12 visits; values > 12 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Mean = 4.01 visits.
  + APO: Mean = 4.28 visits.
  + Difference: +0.27 visits (APO women have more visits).
* **Interpretation**:
  + APO women have slightly more antenatal visits (4.28 vs. 4.01), which is unexpected since more visits are typically associated with better outcomes. This might reflect higher-risk pregnancies in the APO group (e.g., history of APO leading to more monitoring) or better access to care (APO women are more urban, wealthier, and educated).
  + **Context with 2022 Kenya DHS**: The DHS reports that 58% of women had 4 or more antenatal visits for their most recent birth, aligning with our averages (4.01–4.28). The slightly higher number in the APO group might indicate more intensive care due to perceived risk.

## 3. Proportion of Cesarean Births (prop\_cesarean)

* **Description**: Proportion of births delivered by cesarean section across all pregnancies (binary: 0 = No, 1 = Yes).
* **Descriptive Statistics**:
  + NonAPO: Proportion = 13.0%.
  + APO: Proportion = 16.3%.
  + Difference: +3.3 percentage points (APO women have more cesareans).
* **Interpretation**:
  + APO women have a higher proportion of cesarean deliveries (16.3% vs. 13.0%), which aligns with expectations: cesareans are often performed in high-risk pregnancies (e.g., preterm birth, fetal distress), which are components of APO.
  + **Context with 2022 Kenya DHS**: The DHS reports a cesarean section rate of 14% for the most recent birth, consistent with our overall rate (13.0%–16.3%). The higher rate in the APO group supports the association between cesareans and adverse outcomes.

## 4. Place of Residence (place\_of\_residence)

* **Description**: Urban or rural residence (binary: Urban, Rural).
* **Descriptive Statistics**:
  + NonAPO: APO rate = 15.93% (urban women).
  + APO: APO rate = 18.39% (urban women).
  + Difference: Urban women are more likely to have APO (+2.46 percentage points).
* **Interpretation**:
  + Urban women have a higher APO rate (18.39% vs. 15.93%), which is unexpected since urban residence is typically associated with better access to healthcare and lower APO risk. This might be due to urban women having more pregnancies (as seen with total\_pregnancies) or higher detection rates of APO due to better healthcare access.
  + **Context with 2022 Kenya DHS**: The DHS reports better health outcomes in urban areas (e.g., 71% of urban women had 4+ antenatal visits vs. 52% rural), but our data suggests that urban women’s higher APO rate might be linked to other factors like lifestyle or stress.

## 5. Wealth Index (wealth\_index)

* **Description**: Household wealth quintile (Poorest, Poorer, Middle, Richer, Richest).
* **Descriptive Statistics**:
  + APO rate in Richest quintile: 18.55%.
  + Difference: Wealthier women (Richest) have a higher APO rate compared to the overall sample (16.3%–18.39%).
* **Interpretation**:
  + Wealthier women have a higher APO rate, which is unexpected since higher wealth is typically associated with better outcomes. This might be confounded by urban residence (wealthier women are more urban) or higher detection rates of APO due to better access to care.
  + **Context with 2022 Kenya DHS**: The DHS shows that wealthier women have better access to care (e.g., 78% of Richest women had 4+ antenatal visits vs. 40% Poorest), but our data suggests that wealth might not protect against APO, possibly due to lifestyle factors or more pregnancies.

## 6. Highest Education Level (highest\_education\_level)

* **Description**: Highest level of education (No Education, Primary, Secondary, Higher).
* **Descriptive Statistics**:
  + APO rate in Higher education: 18.51%.
  + Difference: Women with higher education have a higher APO rate compared to the overall sample (16.3%–18.39%).
* **Interpretation**:
  + Women with higher education have a higher APO rate, which is unexpected since education is typically protective. This might be confounded by urban residence and wealth (educated women are more urban and wealthier) or higher detection rates of APO.
  + **Context with 2022 Kenya DHS**: The DHS shows that educated women have better health outcomes (e.g., 76% of women with secondary education had 4+ antenatal visits vs. 38% with no education), but our data suggests that education might not protect against APO, possibly due to confounding factors.

## 7. Average Tetanus Vaccinations During Pregnancy (avg\_tetanus\_during\_pregnancy)

* **Description**: Average number of tetanus vaccinations received during pregnancy (capped at 5; values > 5 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Mean = 1.50 vaccinations.
  + APO: Mean = 1.81 vaccinations.
  + Difference: +0.31 vaccinations (APO women have more vaccinations).
* **Interpretation**:
  + APO women have more tetanus vaccinations during pregnancy (1.81 vs. 1.50), which might reflect higher-risk pregnancies (e.g., history of APO prompting more vaccinations) or better access to care (APO women are more urban, wealthier, and educated).
  + **Context with 2022 Kenya DHS**: The DHS reports that 70% of women received 2+ tetanus doses during their last pregnancy, but our averages (1.50–1.81) are lower, possibly because we’re averaging across all pregnancies, including those with no vaccinations.

## 8. Average Succeeding Birth Interval (avg\_succeeding\_birth\_interval)

* **Description**: Average interval between consecutive births (in months, capped at 120 months; values > 120 or < 0 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Mean = 41.7 months.
  + APO: Mean = 43.2 months.
  + Difference: +1.5 months (APO women have longer intervals).
* **Interpretation**:
  + APO women have slightly longer birth intervals (43.2 vs. 41.7 months), which is unexpected since shorter intervals (<24 months) are typically associated with higher APO risk. This might reflect that APO women, being more educated and urban, are better able to space births, but it doesn’t seem to protect against APO.
  + **Context with 2022 Kenya DHS**: The DHS reports a median birth interval of 39 months, slightly lower than our averages (41.7–43.2), possibly because our sample includes all intervals, not just recent ones.

## 9. Proportion of Short Succeeding Intervals (prop\_short\_succeeding\_interval)

* **Description**: Proportion of birth intervals less than 24 months.
* **Descriptive Statistics**:
  + NonAPO: Proportion = 23.2%.
  + APO: Proportion = 21.8%.
  + Difference: -1.4 percentage points (APO women have fewer short intervals).
* **Interpretation**:
  + APO women have a lower proportion of short birth intervals (21.8% vs. 23.2%), which is unexpected since short intervals are a known risk factor for APO. This might be due to APO women being more educated and urban, leading to better birth spacing practices.
  + **Context with 2022 Kenya DHS**: The DHS doesn’t directly report short intervals, but the median interval of 39 months suggests that short intervals (<24 months) are less common, aligning with our data (21.8%–23.2%).

## 10. Average Tetanus Vaccinations Before Birth (avg\_tetanus\_before\_birth)

* **Description**: Average number of tetanus vaccinations received before pregnancy (capped at 5; values > 5 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Mean = 1.27 vaccinations.
  + APO: Mean = 1.10 vaccinations.
  + Difference: -0.17 vaccinations (APO women have fewer vaccinations).
* **Interpretation**:
  + APO women have fewer tetanus vaccinations before pregnancy (1.10 vs. 1.27), which might indicate lower pre-pregnancy healthcare access or younger age at first birth (less time to receive vaccinations). This could contribute to higher APO risk.
  + **Context with 2022 Kenya DHS**: The DHS focuses on tetanus vaccinations during pregnancy, not before, so direct comparison is limited. However, lower pre-pregnancy vaccinations in the APO group might reflect missed opportunities for protection.

## 11. Age at First Birth (Categorized) (age\_at\_first\_birth\_cat)

* **Description**: Age at first birth categorized into three groups: Below 19, 19-35, Above 35 (values < 10 or > 50 recoded to NA).
* **Descriptive Statistics**:
  + NonAPO: Below 19: 39.7%, 19-35: 60.1%, Above 35: 0.22%.
  + APO: Below 19: 35.6%, 19-35: 59.4%, Above 35: 0.23%, NA: 4.74%.
  + Difference: NonAPO women have a higher proportion in "Below 19" (+4.1 percentage points).
* **Interpretation**:
  + Counterintuitively, NonAPO women have a higher proportion of first births before age 19 (39.7% vs. 35.6%), despite younger age at first birth being a known risk factor for APO. This might be due to confounding (e.g., rural NonAPO women having earlier first births) or the small difference (4.1%) not being practically significant.
  + **Context with 2022 Kenya DHS**: The DHS reports a teenage pregnancy rate of 15% (down from 18% in 2014), but our data shows a higher proportion of first births before 19 (35.6%–39.7%), likely because our sample focuses on women with pregnancy outcomes, excluding those who have never been pregnant.

## 12. Abused During Pregnancy (abused\_during\_pregnancy)

* **Description**: Whether the woman experienced abuse during pregnancy (binary: No, Yes; aggregated from multiple pregnancies).
* **Descriptive Statistics**:
  + NonAPO: No: 3.26%, Yes: 55.1%, NA: 41.6%.
  + APO: No: 5.27%, Yes: 51.0%, NA: 43.8%.
  + Difference: NonAPO women report more abuse (+4.1 percentage points).
* **Interpretation**:
  + Counterintuitively, NonAPO women report more abuse during pregnancy (55.1% vs. 51.0%), despite abuse being a known risk factor for APO. The high proportion of missing data (41.6%–43.8%) might bias results, as APO women might underreport abuse due to trauma or stigma.
  + **Context with 2022 Kenya DHS**: The DHS reports that 34% of women experienced physical violence since age 15, but our data shows higher rates (51.0%–55.1%), possibly because we’re capturing abuse specifically during pregnancy, which might be more prevalent or better recalled.

## 13. Average Size of Child (avg\_size\_of\_child)

* **Description**: Average perceived size of children across all births (1 = Very Small, 5 = Very Large; values outside 1–5 recoded to NA).
* **Descriptive Statistics**:
  + Summary: Min = 1, 1st Qu. = 3, Median = 3, Mean = 2.971, 3rd Qu. = 3, Max = 5, NA’s = 18,159 (77.9% of 23,317 women).
  + NonAPO: Mean = 2.96.
  + APO: Mean = 3.02.
  + Difference: +0.06 (APO women have slightly larger children).
* **Interpretation**:
  + Counterintuitively, APO women have a slightly higher average child size (3.02 vs. 2.96), despite smaller child size (e.g., low birth weight) being a component of APO. The difference (0.06) is very small and likely not practically significant. The high proportion of missing data (77.9%) limits reliability, as missingness might be non-random (e.g., rural women less likely to report).
  + **Context with 2022 Kenya DHS**: The DHS reports that 7% of births are low birth weight (<2.5 kg), but our data uses perceived size (1–5 scale), not weight, and averages across all births, which might dilute the effect of low birth weight in APO cases.

## 14. Total Pregnancies (total\_pregnancies)

* **Description**: Total number of pregnancies (including live births, stillbirths, miscarriages, abortions).
* **Descriptive Statistics**:
  + Summary: Min = 1, 1st Qu. = 2, Median = 3, Mean = 3.534, 3rd Qu. = 5, Max = 16.
  + NonAPO: Mean = 3.26 pregnancies.
  + APO: Mean = 4.87 pregnancies.
  + Difference: +1.61 pregnancies (APO women have more pregnancies).
* **Interpretation**:
  + APO women have more pregnancies (4.87 vs. 3.26), which aligns with expectations: each pregnancy is an opportunity for an adverse outcome, increasing APO risk. The difference (1.61) is substantial and suggests that total\_pregnancies is a strong predictor of APO.
  + **Context with 2022 Kenya DHS**: The DHS reports a total fertility rate of 3.4 children per woman, but our data includes all pregnancies (not just live births), explaining the higher mean (3.534). The higher number in the APO group supports the exposure hypothesis.

## 15. Current Contraceptive Method (contraceptive\_category)

* **Description**: Current contraceptive method categorized into None, Modern, Traditional, Other (based on 2022 Kenya DHS coding).
* **Descriptive Statistics**:
  + NonAPO: None: 46.3%, Modern: 48.7%, Traditional: 4.96%.
  + APO: None: 51.9%, Modern: 41.5%, Traditional: 6.59%.
  + Difference: APO women have higher non-use (+5.6 percentage points) and lower modern method use (-7.2 percentage points).
* **Interpretation**:
  + APO women are more likely to use no method (51.9% vs. 46.3%) and less likely to use modern methods (41.5% vs. 48.7%), which aligns with expectations: lack of effective contraception can lead to unintended pregnancies, increasing APO risk. The slightly higher use of traditional methods in the APO group (6.59% vs. 4.96%) might also contribute, as traditional methods are less effective.
  + **Context with 2022 Kenya DHS**: The DHS reports that among married women, 57% use modern methods, 5% use traditional methods, and 38% use no method. Our sample shows lower modern method use (41.5%–48.7%) and higher non-use (46.3%–51.9%), likely because it includes all women (not just married) and reflects higher unmet need (14% among married women per DHS).

## 16. Respondent Circumcised (respondent\_circumcised)

* **Description**: Whether the respondent has undergone female circumcision (FGM/C) (binary: No, Yes).
* **Descriptive Statistics**:
  + NonAPO: NA: 100%.
  + APO: NA: 100%.
* **Interpretation**:
  + The variable is entirely missing (100% NA), so it cannot be analyzed or included in the model. This is a limitation, as FGM/C is a known risk factor for APO in Kenya, where prevalence is 21% among women aged 15–49 (2022 Kenya DHS). The absence of data might be due to non-collection in the survey or high non-response due to sensitivity.
  + **Context with 2022 Kenya DHS**: The DHS reports that 21% of women have undergone FGM/C, with higher prevalence in rural areas (24% vs. 15% urban). Our inability to analyze this variable means we miss a potential predictor of APO.

## Overall Discussion

* **Expected Patterns**:
  + Predictors like total\_pregnancies (more pregnancies in APO), prop\_cesarean (higher in APO), and contraceptive\_category (higher non-use in APO) align with expectations and are likely significant in a logistic regression model.
  + Other predictors like avg\_timing\_first\_antenatal\_check, avg\_num\_antenatal\_visits, place\_of\_residence, wealth\_index, highest\_education\_level, and avg\_tetanus\_during\_pregnancy show patterns that are partially expected but complicated by confounding (e.g., APO women have better access to care but higher APO rates).
* **Counterintuitive Patterns**:
  + age\_at\_first\_birth\_cat, abused\_during\_pregnancy, and avg\_size\_of\_child show counterintuitive results:
    - Younger age at first birth (Below 19) is more common in NonAPO, not APO.
    - Abuse during pregnancy is more common in NonAPO, not APO.
    - Average child size is slightly higher in APO, not lower.
  + These findings might be due to confounding (e.g., rural NonAPO women having earlier first births and more abuse) or high missing data (abused\_during\_pregnancy: 41.6%–43.8%, avg\_size\_of\_child: 77.9%) introducing bias.
* **Context with 2022 Kenya DHS**:
  + The DHS highlights challenges like unmet need for family planning (14%), teenage pregnancy (15%), and violence (34% of women), which align with some of our findings (e.g., high non-use of contraception, high rates of abuse). However, our sample’s focus on women with pregnancy outcomes and inclusion of all pregnancies (not just recent ones) explains some discrepancies (e.g., higher teenage first births, lower modern method use).
* **Implications for Modeling**:
  + All predictors except respondent\_circumcised should be included in a logistic regression model to assess their combined effects on apo\_general.
  + High missing data in abused\_during\_pregnancy and avg\_size\_of\_child might reduce model performance, and imputation or sensitivity analyses may be needed.
  + Interactions (e.g., contraceptive\_category and wealth\_index) could be explored to understand how access to contraception varies by socioeconomic status.